



US009475856B2

(12) **United States Patent**
Mohammadi et al.

(10) **Patent No.:** US 9,475,856 B2
(b4) **Date of Patent:** Oct. 25, 2016

(54) **CHIMERIC FGF21 PROTEINS WITH ENHANCED BINDING AFFINITY FOR β -KLOTHO FOR THE TREATMENT OF TYPE II DIABETES, OBESITY, AND RELATED METABOLIC DISORDERS**

(71) Applicant: **New York University**, New York, NY (US)

(72) Inventors: **Moosa Mohammadi**, Scarsdale, NY (US); **Regina Goetz**, New York, NY (US)

(73) Assignee: **New York University**, New York, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 157 days.

(21) Appl. No.: **13/784,289**

(22) Filed: **Mar. 4, 2013**

(65) **Prior Publication Data**

US 2013/0231277 A1 Sep. 5, 2013

Related U.S. Application Data

(60) Provisional application No. 61/605,961, filed on Mar. 2, 2012.

(51) **Int. Cl.**

A61K 38/18 (2006.01)
C07K 14/50 (2006.01)
A61K 45/06 (2006.01)
G01N 33/74 (2006.01)
A61K 38/00 (2006.01)

(52) **U.S. Cl.**

CPC **C07K 14/50** (2013.01); **A61K 38/1825** (2013.01); **A61K 45/06** (2013.01); **G01N 33/74** (2013.01); **A61K 38/00** (2013.01); **C07K 2319/00** (2013.01); **C07K 2319/21** (2013.01); **G01N 2333/50** (2013.01); **G01N 2500/02** (2013.01); **G01N 2800/042** (2013.01)

(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,132,408 A	7/1992	Baird et al.
5,478,804 A	12/1995	Calabresi et al.
6,326,484 B1	12/2001	Gage et al.
6,982,170 B1	1/2006	Maciag et al.
7,491,697 B2	2/2009	Beals et al.
7,582,607 B2	9/2009	Frye et al.
7,622,445 B2	11/2009	Frye et al.
7,655,627 B2	2/2010	Frye et al.
7,956,033 B2	6/2011	Cheng et al.
8,168,591 B2	5/2012	Takada et al.
8,642,546 B2 *	2/2014	Belouski et al. 514/9.1
8,889,426 B2	11/2014	Mohammadi et al.
8,889,621 B2	11/2014	Mohammadi et al.
8,906,854 B2	12/2014	Jonker et al.
8,951,966 B2 *	2/2015	Ling et al. 514/9.1
8,999,929 B2	4/2015	Mohammadi et al.
9,072,708 B2	7/2015	Jonker et al.
9,272,017 B2	3/2016	Mohammadi et al.
2004/0259780 A1	12/2004	Glasebrook et al.

2007/0142278 A1	6/2007	Beals et al.
2007/0237768 A1	10/2007	Glaesner et al.
2007/0265200 A1	11/2007	Glaesner et al.
2007/0293430 A1	12/2007	Frye et al.
2007/0299007 A1	12/2007	Frye et al.
2008/0103096 A1	5/2008	Frye et al.
2008/0255045 A1	10/2008	Cujec et al.
2008/0261875 A1	10/2008	Etgen et al.
2009/0111742 A1	4/2009	Kharitonenkova et al.
2009/0118190 A1	5/2009	Beals et al.
2009/0305986 A1	12/2009	Belouski et al.
2010/0062984 A1	3/2010	Kumar et al.
2010/0158914 A1	6/2010	Desnoyers
2010/0184665 A1	7/2010	Suzuki et al.
2010/0216715 A1	8/2010	Tagmose et al.
2010/0285131 A1	11/2010	Belouski et al.
2010/0286042 A1	11/2010	Imamura et al.

(Continued)

FOREIGN PATENT DOCUMENTS

EP	0 645 451 B1	8/2001
WO	2011/047267 A1	4/2011

(Continued)

OTHER PUBLICATIONS

Presta et al., "Structure-Function Relationship of Basic Fibroblast Growth Factor: Site-Directed Mutagenesis of a Putative Heparin-Binding and Receptor-Binding Region," *Biochem. Biophys. Res. Commun.* 185(3):1098-1107 (1992).

Zakrzewska et al., "Increased Protein Stability of FGF1 Can Compensate for Its Reduced Affinity for Heparin," *J. Biol. Chem.* 284(37):25388-403 (2009).

Motomura et al., "An FGF1:FGF2 Chimeric Growth Factor Exhibits Universal FGF Receptor Specificity, Enhanced Stability and Augmented Activity Useful for Epithelial Proliferation and Radioprotection," *Biochim. Biophys. Acta* 1780 (12):1432-40 (2008).

(Continued)

Primary Examiner — Christine J Saoud

(74) *Attorney, Agent, or Firm* — LeClairRyan, a Professional Corporation

(57)

ABSTRACT

The present invention relates to chimeric proteins that include an N-terminus coupled to a C-terminus, where the N-terminus includes an N-terminal portion of fibroblast growth factor 21 ("FGF21") and the C-terminus includes a C-terminal portion of fibroblast growth factor 19 ("FGF19"). The present invention also relates to pharmaceutical compositions including chimeric proteins according to the present invention, as well as methods for treating a subject suffering from diabetes, obesity, or metabolic syndrome, methods of treating a subject in need of increased FGF21- β Klotho-FGF receptor complex formation, methods of causing increased FGF21 receptor agonist- β Klotho-FGF receptor complex formation, and methods of screening for compounds with enhanced binding affinity for the β Klotho-FGF receptor complex involving the use of chimeric proteins of the present invention.

(56)

References Cited**U.S. PATENT DOCUMENTS**

- 2010/0323954 A1 12/2010 Li et al.
 2011/0053841 A1 3/2011 Yayon et al.
 2011/0104152 A1 5/2011 Sonoda
 2011/0150901 A1 6/2011 Smith et al.
 2011/0172401 A1 7/2011 Cujec et al.
 2011/0190207 A1 8/2011 Mohammadi et al.
 2011/0195077 A1 8/2011 Glass et al.
 2012/0052069 A1 3/2012 Belouski et al.
 2012/0288886 A1 11/2012 Mohammadi et al.
 2013/0023474 A1 1/2013 Ling et al.
 2013/0058896 A1 3/2013 Takada et al.
 2013/0116171 A1 5/2013 Jonker et al.
 2013/0184211 A1 7/2013 Mohammadi et al.
 2013/0331316 A1 12/2013 Mohammadi et al.
 2013/0331317 A1 12/2013 Mohammadi et al.
 2013/0331325 A1 12/2013 Mohammadi et al.
 2014/0094406 A1 4/2014 Mohammadi et al.
 2014/0107022 A1 4/2014 Mohammadi et al.
 2014/0155316 A1 6/2014 Mohammadi et al.
 2014/0171361 A1 6/2014 Jonker et al.
 2014/0243260 A1 8/2014 Mohammadi et al.
 2015/0111821 A1 4/2015 Suh et al.
 2015/0343022 A1 12/2015 Jonker et al.

FOREIGN PATENT DOCUMENTS

- WO 2011/130729 A2 10/2011
 WO 2013/184958 A1 12/2013
 WO 2013/184960 A2 12/2013
 WO 2013/184962 A1 12/2013
 WO 2015/149069 A1 10/2015

OTHER PUBLICATIONS

- Nakayama et al., "Post Treatment With an FGF Chimeric Growth Factor Enhances Epithelial Cell Proliferation to Improve Recovery From Radiation-Induced Intestinal Damage," *Int. J. Radiat. Oncol. Biol. Phys.* 78(3):860-7 (2010).
- Kharitonenkov et al., "The Metabolic State of Diabetic Monkeys is Regulated by Fibroblast Growth Factor-21," *Endocrinology* 148(2):774-81 (2007).
- Igarashi et al., "Characterization of Recombinant Human Fibroblast Growth Factor (FGF)-10 Reveals Functional Similarities With Keratinocyte Growth Factor (FGF-7)," *J. Biol. Chem.* 273(21):13230-5 (1998).
- Goetz et al., "Isolated C-Terminal tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *PNAS* 107(1):407-412 (Epub Dec. 4, 2009).
- Goetz et al., "Molecular Insights Into the Klotho-Dependent, Endocrine Mode of Action of Fibroblast Growth Factor 19 Subfamily Members," *Mol. Cell. Biol.* 27(9):3417-3428 (2007).
- Beenken, "Structural and Biochemical Studies of FGF-FGFR Complexes," Thesis (Sep. 2011).
- Ge et al., "Characterization of a FGF19 Variant With Altered Receptor Specificity Revealed a Central Role for FGFR1c in the Regulation of Glucose Metabolism," *PLoS One*, 7(3):e33603 (Epub Mar. 23, 2012).
- Wu et al., "FGF19 Regulates Cell Proliferation, Glucose and Bile Acid Metabolism Via FGFR4-Dependent and Independent Pathways," *PLoS One* 6(3):e17868 (Mar. 8, 2011).
- Wu et al., "Selective Activation of FGFR4 by an FGF19 Variant Does Not Improve Glucose Metabolism in OB/OB Mice," *Proc. Nat'l. Acad. Sci. U.S.A.* 106(34):14379-84 (2009).
- Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine & Growth Factor Reviews* 16:107-137 (2005).
- Hutley et al., "Fibroblast Growth Factor 1: A Key Regulator of Human Adipogenesis," *Diabetes* 53:3097-3106 (2004).
- Imamura et al., "Recovery of Mitogenic Activity of a Growth Factor Mutant with Nuclear Translocation Sequence," *Science* 249:1567-1570 (Sep. 28, 1990).
- International Search Report and Written Opinion for PCT/US13/44589 (Nov. 13, 2013).
- International Search Report and Written Opinion for PCT/US13/44594 (Nov. 13, 2013).
- International Search Report and Written Opinion for PCT/US13/44592 (Jan. 17, 2014).
- International Search Report and Written Opinion for corresponding PCT/US2013/028888 (Jul. 23, 2013).
- International Search Report and Written Opinion for PCT/US14/17367 (Jun. 18, 2014).
- Razzaque, "The FGF23-Klotho Axis: Endocrine Regulation of Phosphate Homeostasis," *Nat. Rev. Endocrinol.* 5 (11):611-19 (2009).
- Yie et al., "FGF21 N- and C-Termini Play Different Roles in Receptor Interaction and Activation," *FEBS Lett.* 583:19-24(2009).
- Andrukova et al., "FGF23 Acts Directly on Renal Proximal Tubules to Induce Phosphaturia Through Activation of the ERK1/2-SKG1 Signaling Pathway," *Bone* 51(3):621-8 (Jun. 12, 2012).
- Beenken et al., "Plasticity in Interactions of Fibroblast Growth Factor 1 (FGF1) N Terminus With FGF Receptors Underlies Promiscuity of FGF1," *J. Biol. Chem.* 287(5):3067-3078 (Nov. 4, 2011).
- Jonker et al., "A PPARGamma-FGF1 Axis Is Required for Adaptive Adipose Remodelling and Metabolic Homeostasis," *Nature* 485(7398):391-394 (Apr. 22, 2012).
- Wu et al., "A Unique FGF23 With the Ability to Activate FGFR Signaling Through Both alphaKlotho and betaKlotho," *J Mol. Biol.* 418:82-89 (2012).
- Beenken & Mohammadi, "The Structural Biology of the FGF19 Subfamily," *Adv. Exp. Med. Biol.* 728:1-24 (2012).
- Wu et al., "C-Terminal Tail of FGF19 Determines Its Specificity Toward Klotho Co-Receptors," *J. Biol. Chem.* 283 (48):33304-33309 (2008).
- Goetz et al., "Conversion of a Paracrine Fibroblast Growth Factor Into an Endocrine Fibroblast Growth Factor," *J. Biol. Chem.* 287(34):29134-29146 (Jun. 25, 2012).
- Goetz et al., "Klotho Coreceptors Inhibit Signaling by Paracrine Fibroblast Growth Factor 8 Subfamily Ligands," *Mol. Cell. Biol.* 32(10):1944-1954 (Mar. 26, 2012).
- Olsen et al., "Insights Into the Molecular Basis for Fibroblast Growth Factor Receptor Autoinhibition and Ligand-Binding Promiscuity," *Proc. Nat'l. Acad. Sci. USA* 101(4):935-940 (2004).
- Wei et al., "Fibroblast Growth Factor 21 Promotes Bone Loss by Potentiating the Effects of Peroxisome Proliferator-Activated Receptor Gamma," *Proc. Nat'l. Acad. Sci. USA* 109(8):3143-3148 (Feb. 21, 2012).
- Wu et al., "Separating Mitogenic and Metabolic Activities of Fibroblast Growth Factor 19 (FGF19)," *Proc. Nat'l. Acad. Sci. USA* 107(32):14158-14163 (2010).
- Wu et al., "FGF19-Induced Hepatocyte Proliferation Is Mediated Through FGFR4 Activation," *J. Biol. Chem.* 285(8):5165-5170 (2009).
- Zhang et al., "Receptor Specificity of the Fibroblast Growth Factor Family," *J. Biol. Chem.* 281(23):15694-15700 (2006).
- PCT Search Report and Written Opinion for PCT/US2013/028888 (Jul. 23, 2013).
- Abraham et al., "Human Basic Fibroblast Growth Factor: Nucleotide Sequence and Genomic Organization," *EMBO J.* 5(10):2523-2528 (1986).
- Esch et al., "Primary Structure of Bovine Pituitary Basic Fibroblast Growth Factor (FGF) and Comparison with the Amino-Terminal Sequence of Bovine Brain Acidic FGF," *PNAS* 82:6507-6511 (1985).
- Kurosu et al., "The Klotho Gene Family as a Regulator of Endocrine Fibroblast Growth Factors," *Mol. Cel. Endocrin.* 299:72-78 (2009).
- Ono et al., "Novel Regulation of Fibroblast Growth Factor 2 (FGF2)-Mediated Cell Growth by Polysialic Acid," *J. Biol. Chem.* 287(6):3710-3722 (2012).
- Schlessinger et al., "Crystal Structure of a Ternary FGF-FGFR-Heparin Complex Reveals a Dual Role for Heparin in FGF Binding and Dimerization," *Molecular Cell* 6:743-750 (2000).

(56)

References Cited

OTHER PUBLICATIONS

Thompson et al., "Energetic Characterization of the Basic Fibroblast Growth Factor-Heparin Interaction: Identification of the Heparin Binding Domains," *Biochemistry* 33:3831-3840 (1994).

Suh et al., "Endocrinization of FGF1 Produces a Neomorphic and Potent Insulin Sensitizer," *Author Manuscript, Nature* 513(7518): 436-439 (2014).

Beenken et al., "The FGF Family: Biology, Pathophysiology and Therapy," *Nat Rev Drug Discov.* 8(3):235-53 (Mar. 2009).

Kurosu et al., "Tissue-specific Expression of β Klotho and Fibroblast Growth Factor (FGF) Receptor Isoforms Determines Metabolic Activity of FGF19 and FGF21," *J. Biol. Chem.* 282(37):26687-26695 (2007).

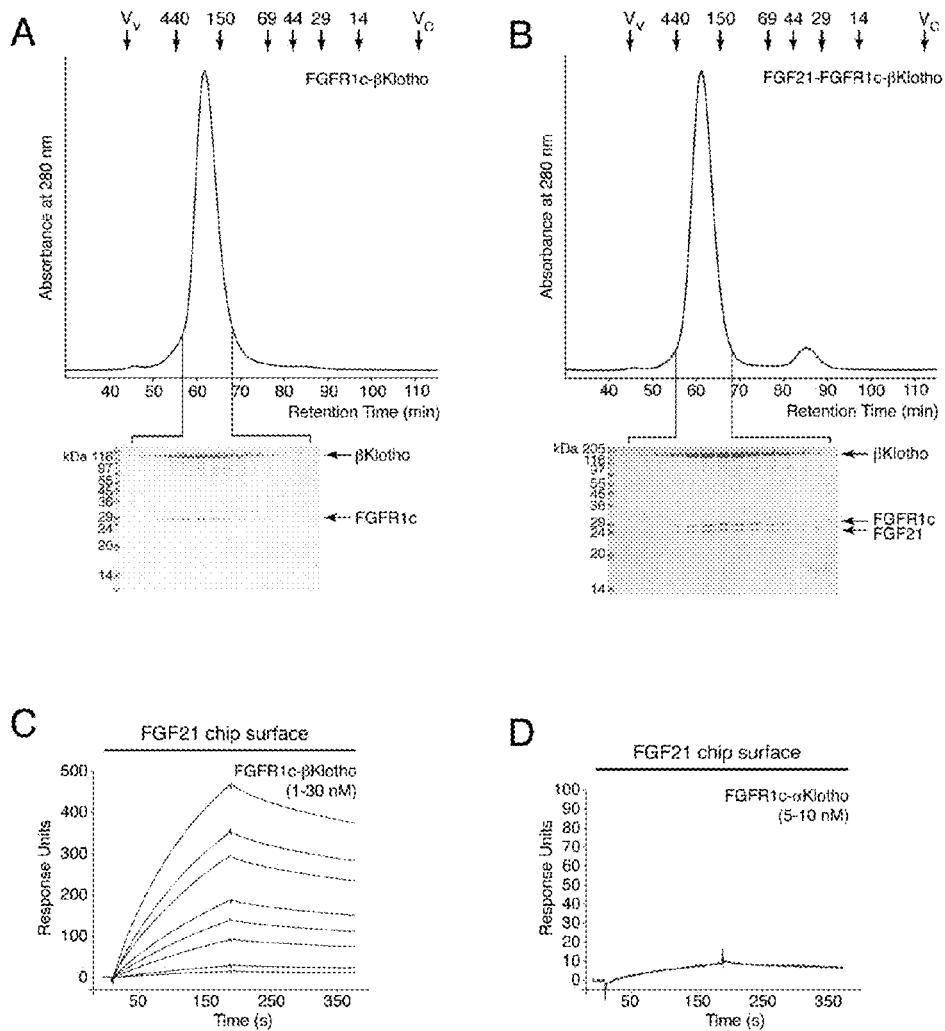
Micanovic et al., "Different Roles of N- and C- Termini in the Functional Activity of FGF21," *J. Cell. Physiol.* 219:227-234 (2009).

Kharitonov et al., "FGF-21/FGF-21 Receptor Interaction and Activation is Determined by β Klotho," *J. Cell. Physiol.* 215:1-7 (2008).

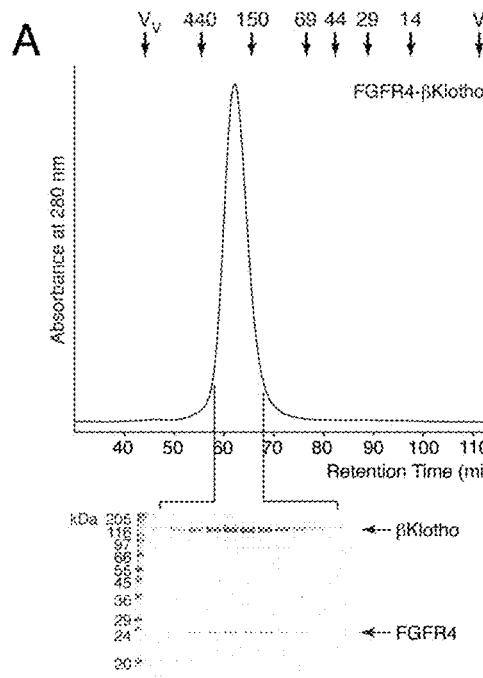
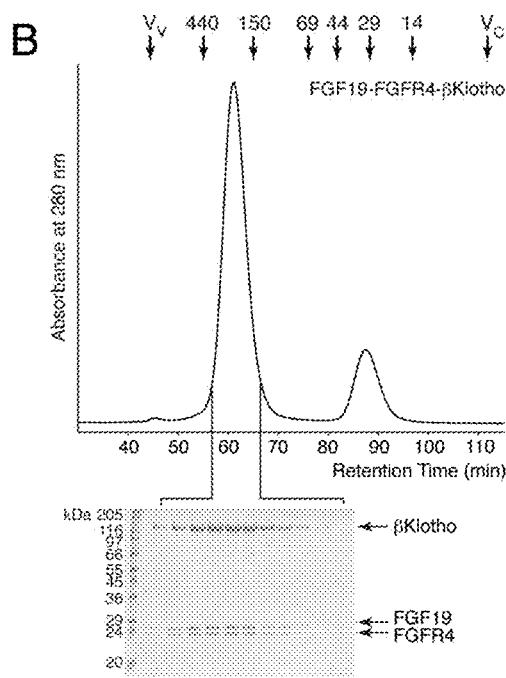
Yao et al., "Expression and Pharmacological Evaluation of Fusion Protein FGF21-L-Fc," *Acta Pharmaceutica Sinica* 46(7):787-92 (2011) (Abstract in English).

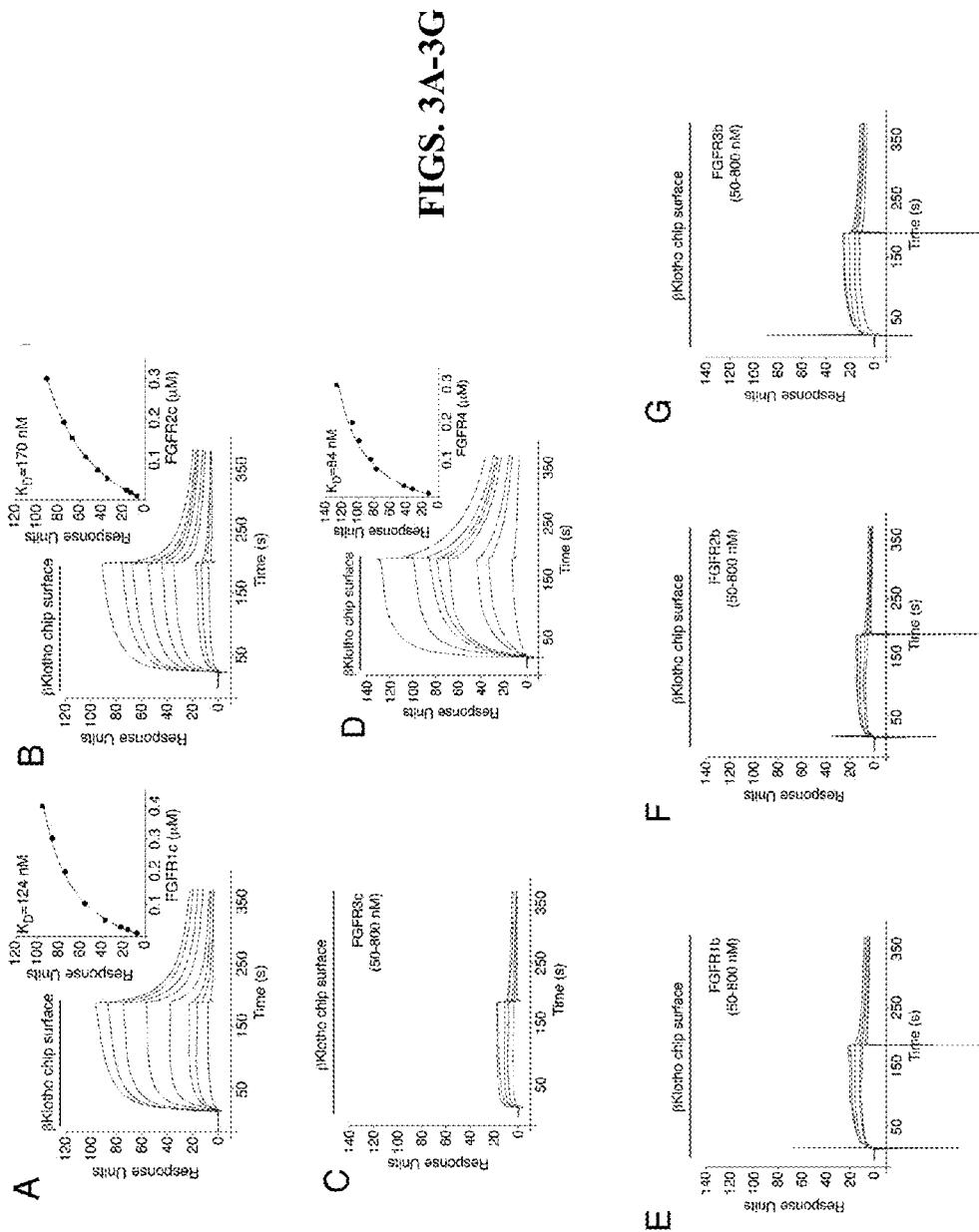
Extended European Search Report for European Application No. 13799858.9, 13 pages (dated May 3, 2016).

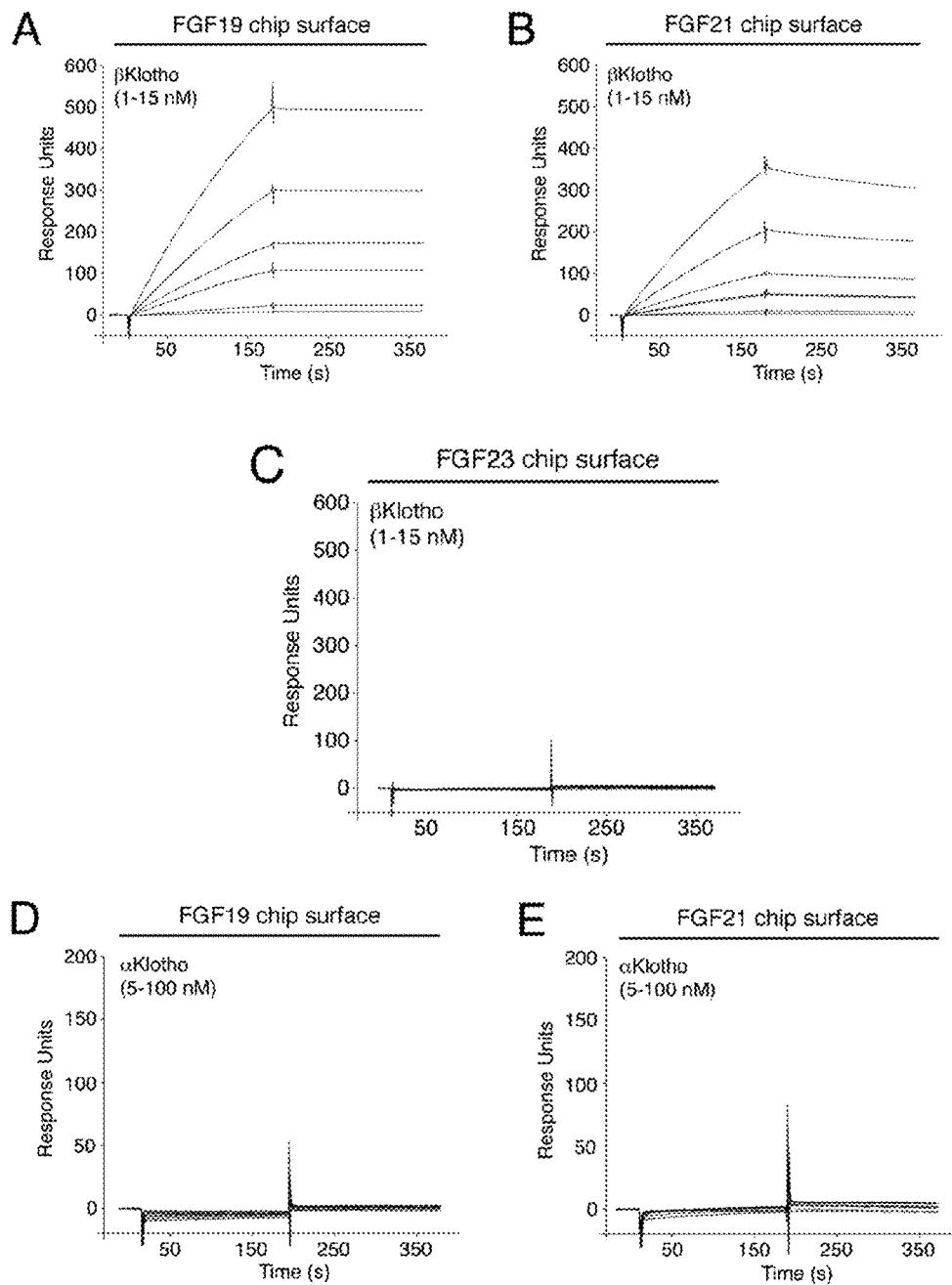
* cited by examiner



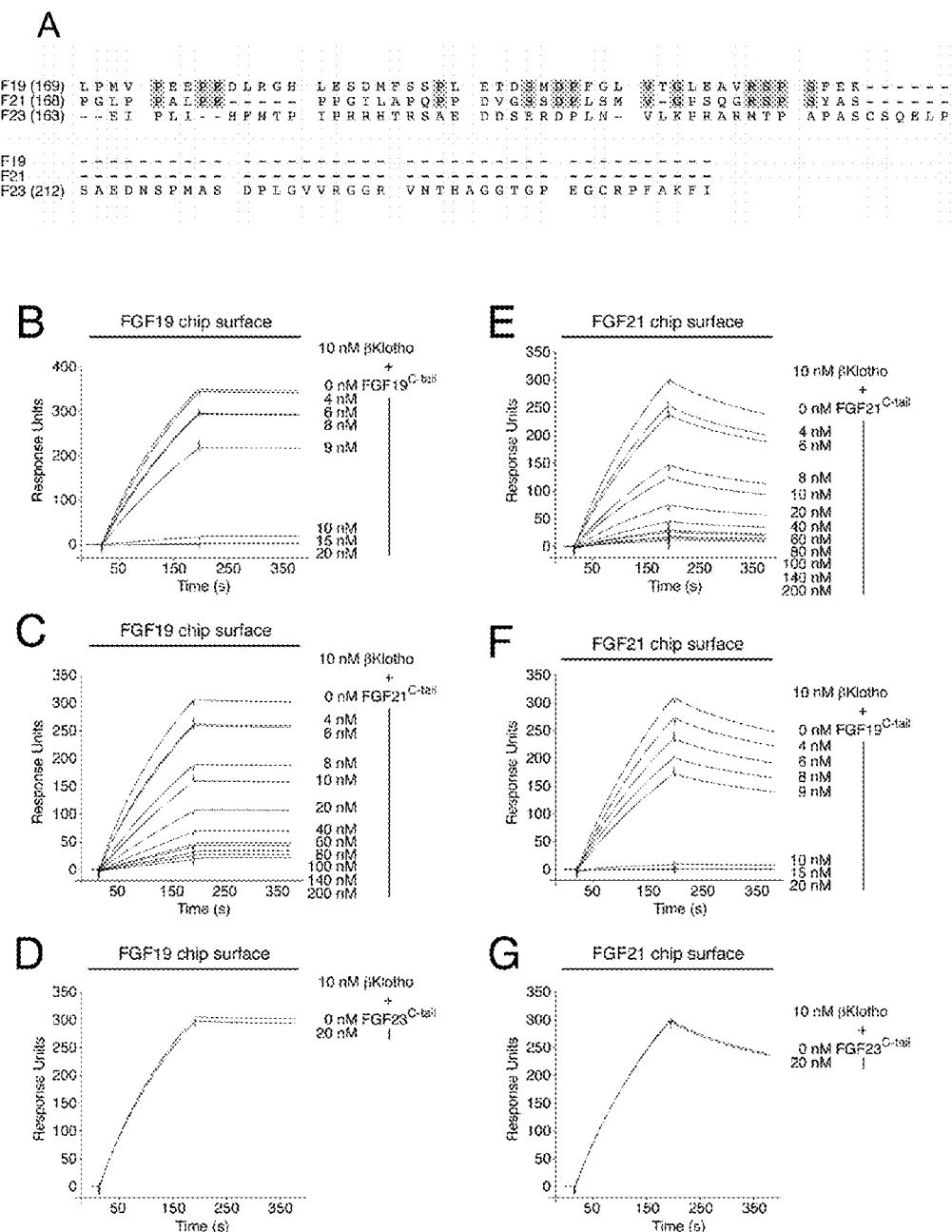
FIGS. 1A-1D

**FIGS. 2A-2B**

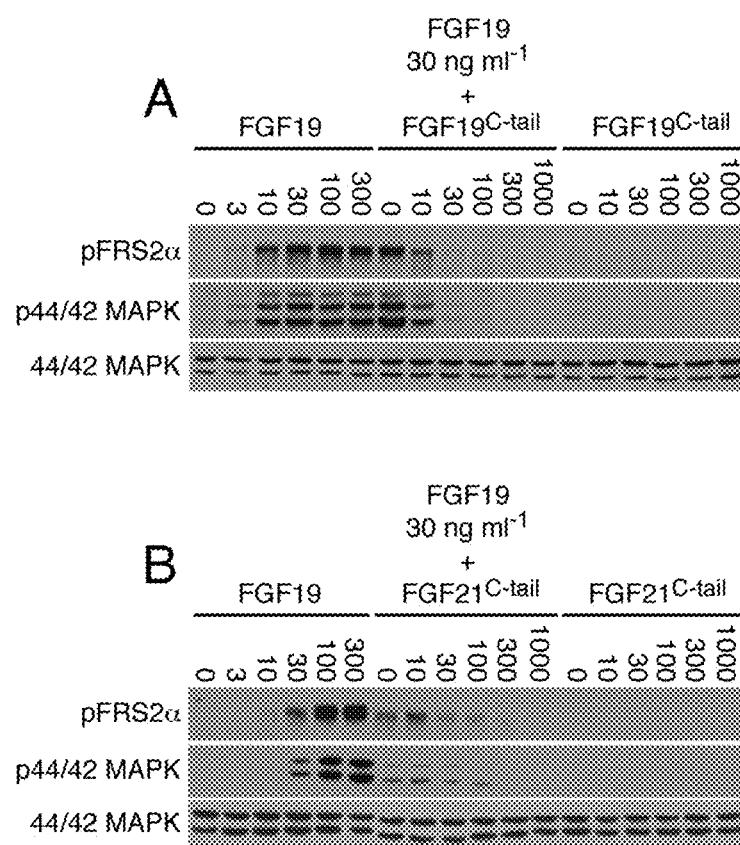


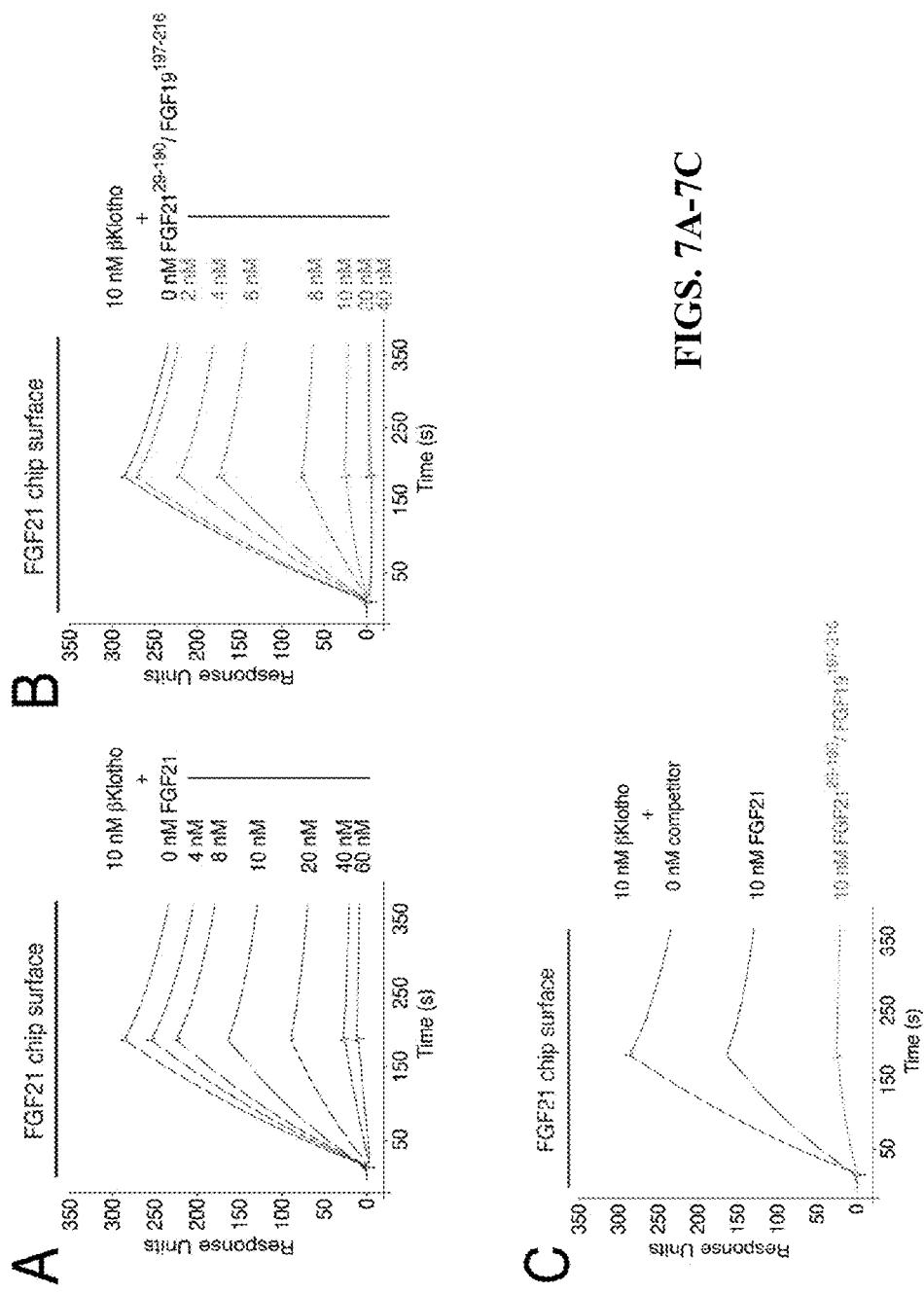


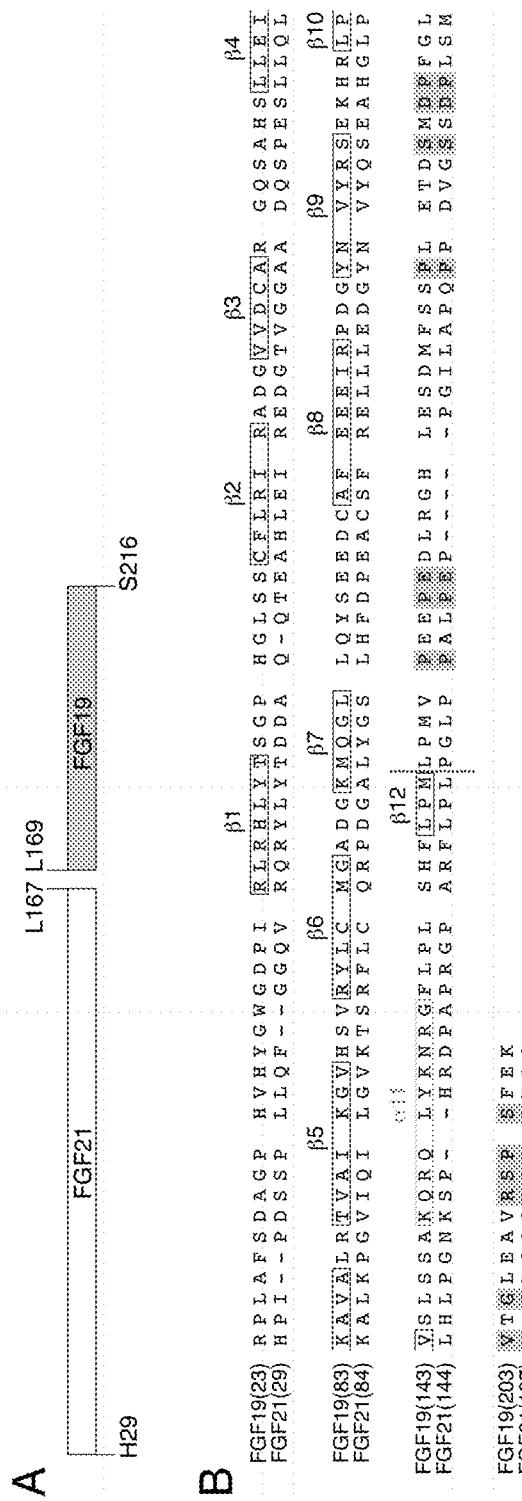
FIGS. 4A-4E



FIGS. 5A-5G

**FIGS. 6A-6B**





FIGS. 8A-8B

FIG. 9

FIG. 9

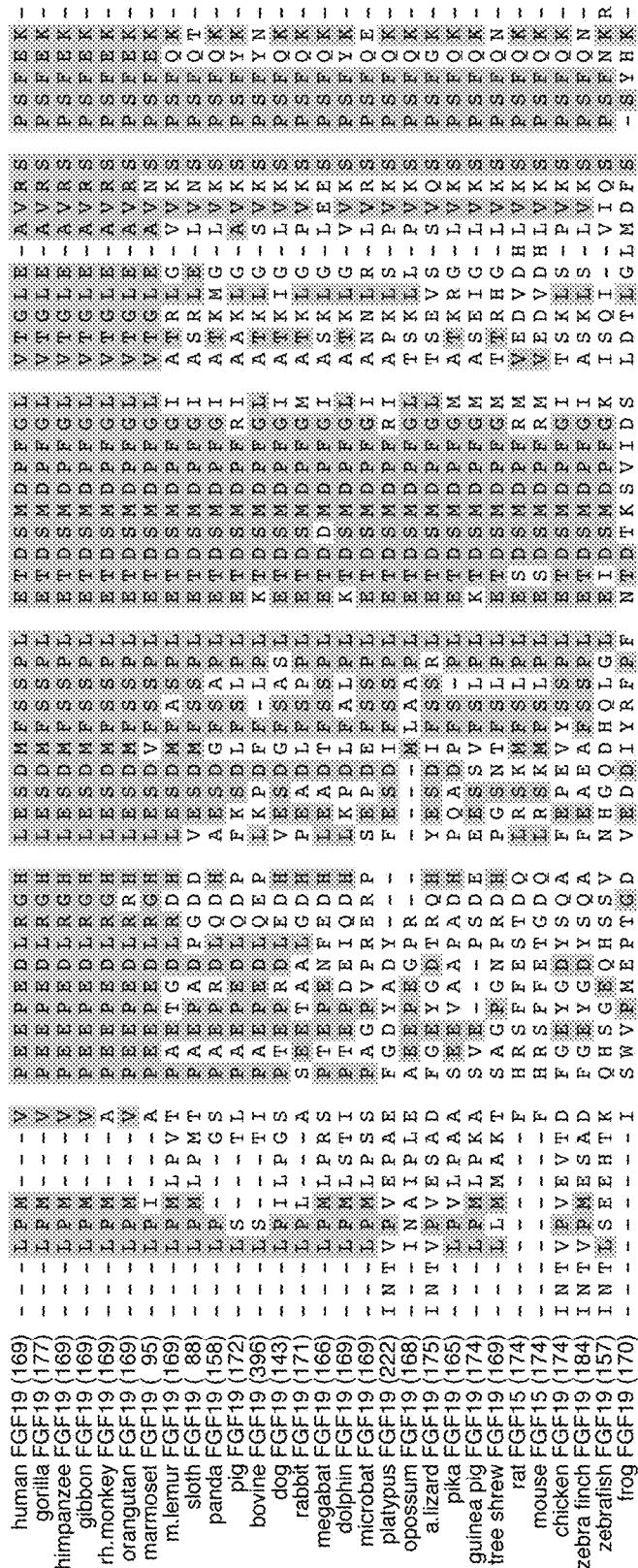


FIG. 10

FGF21C-tail(168) P G L P P A L S K (SEQ ID NO:234)
 FGF21C-tail19-1 P G L P P A L S (SEQ ID NO:235)
 FGF21C-tail19-2 P G L P P A L S (SEQ ID NO:236)
 FGF21C-tail19-3 P G L P P A L S (SEQ ID NO:237)
 FGF21C-tail19-4 P G L P P A L S (SEQ ID NO:238)
 FGF21C-tail19-5 P G L P P A L S (SEQ ID NO:239)
 FGF21C-tail19-6 P G L P P A L S (SEQ ID NO:240)
 FGF21C-tail19-7 P G L P P A L S (SEQ ID NO:241)
 FGF21C-tail19-8 P G L P P A L S (SEQ ID NO:242)
 FGF21C-tail19-9 P G L P P A L S (SEQ ID NO:243)
 FGF21C-tail19-10 P G L P P A L S (SEQ ID NO:244)
 FGF21C-tail19-11 P G L P P A L S (SEQ ID NO:245)
 FGF21C-tail19-12 P G L P P A L S (SEQ ID NO:246)
 FGF21C-tail19-13 P G L P P A L S (SEQ ID NO:247)
 FGF21C-tail19-14 P G L P P A L S (SEQ ID NO:248)
 FGF21C-tail19-15 P G L P P A L S (SEQ ID NO:249)
 FGF21C-tail19-16 P G L P P A L S (SEQ ID NO:250)
 FGF21C-tail19-17 P G L P P A L S (SEQ ID NO:251)
 FGF21C-tail19-18 P G L P P A L S (SEQ ID NO:252)
 FGF21C-tail19-19 P G L P P A L S (SEQ ID NO:253)
 FGF21C-tail19-20 P G L P P A L S (SEQ ID NO:254)
 FGF21C-tail19-21 P G L P P A L S (SEQ ID NO:255)
 FGF21C-tail19-22 P G L P P A L S (SEQ ID NO:256)
 FGF21C-tail19-23 P G L P P A L S (SEQ ID NO:257)
 FGF21C-tail19-24 P G L P P A L S (SEQ ID NO:258)
 FGF21C-tail19-25 P G L P P A L S (SEQ ID NO:259)
 FGF21C-tail19-26 P G L P P A L S (SEQ ID NO:260)
 FGF21C-tail19-27 P G L P P A L S (SEQ ID NO:261)
 FGF21C-tail19-28 P G L P P A L S (SEQ ID NO:262)
 FGF21C-tail19-29 P G L P P A L S (SEQ ID NO:263)
 FGF21C-tail19-30 P G L P P A L S (SEQ ID NO:264)
 FGF21C-tail19-31 P G L P P A L S (SEQ ID NO:265)
 FGF21C-tail19-32 P G L P P A L S (SEQ ID NO:266)
 FGF21C-tail19-33 P G M P P A L S (SEQ ID NO:267)
 FGF21C-tail19-34 P L P P A L S (SEQ ID NO:268)
 FGF19C-tail(169) L P M V P P E E D M F S S P L E T D D M F S S P L E T E D M F D R G H L E S P E D I L R G H L E S LEAV E R S P P F E K

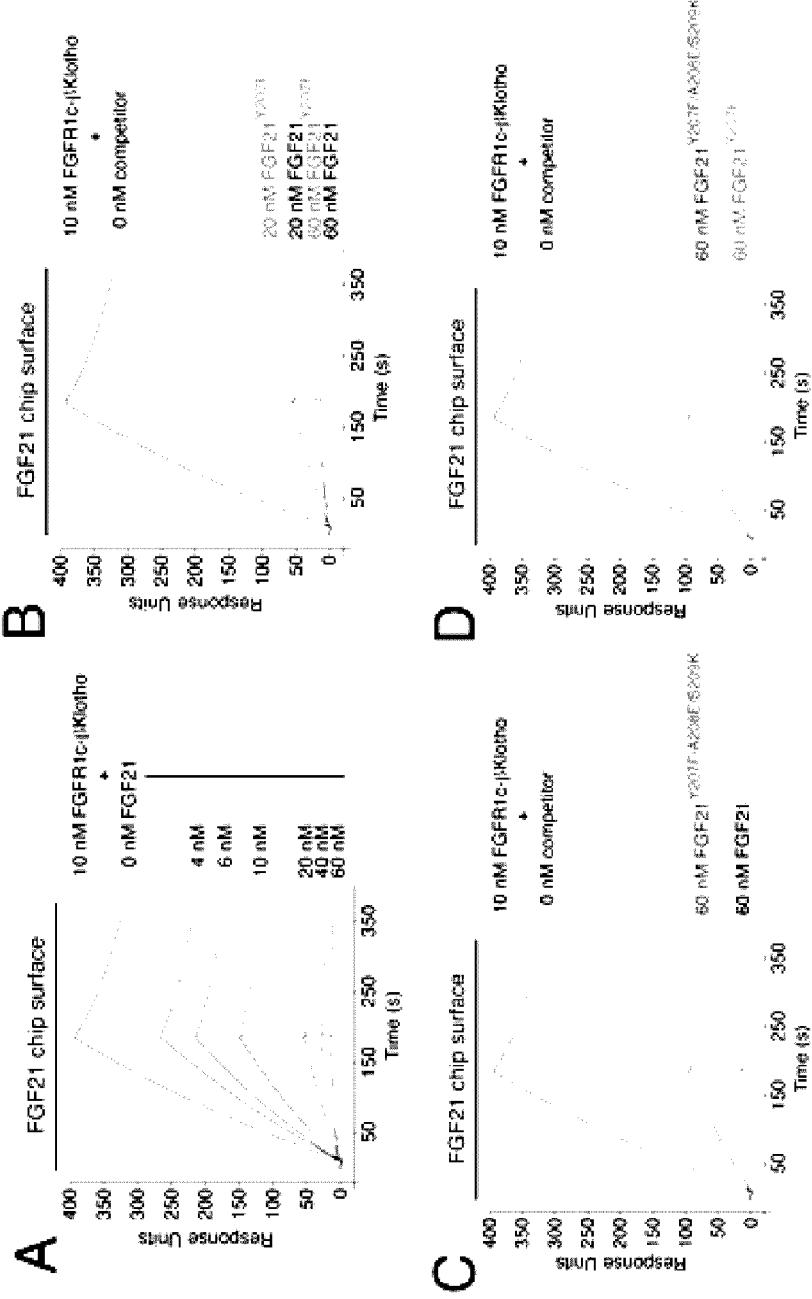
FIG. 11

FGF21C-tail(168)	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S Q G R S P S Y A	K (SEQ ID NO:234)
FGF21C-tail19-1	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S Q G R S P S Y A	K (SEQ ID NO:238)
FGF21C-tail19-35	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S Q G R S P S Y E	K (SEQ ID NO:269)
FGF21C-tail19-36	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S Q G R S P S Y E	K (SEQ ID NO:270)
FGF21C-tail19-37	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S Q V R S P S F E	K (SEQ ID NO:271)
FGF21C-tail19-38	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P S A V R S P S F E	K (SEQ ID NO:272)
FGF21C-tail19-39	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	P E A V R S P S F E	K (SEQ ID NO:273)
FGF21C-tail19-40	P G L P P A I	P E - - - - -	P P G	S S D P L S M V - G	LE A V R S P S F E	K (SEQ ID NO:274)
FGF21C-tail19-41	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:275)
FGF21C-tail19-42	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:276)
FGF21C-tail19-43	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:277)
FGF21C-tail19-44	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:278)
FGF21C-tail19-45	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:279)
FGF21C-tail19-46	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:280)
FGF21C-tail19-47	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:281)
FGF21C-tail19-48	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:282)
FGF21C-tail19-49	P G L P P A I	P E - - - - -	P P G	I L A P Q P P D V G	LE A V R S P S F E	K (SEQ ID NO:283)
FGF21C-tail19-50	P G L P P A I	P E - - - - -	P P G	I L A P S P L E T D	S M D P F G L V T G	K (SEQ ID NO:284)
FGF21C-tail19-51	P G L P P A I	P E - - - - -	P P G	I L A S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:285)
FGF21C-tail19-52	P G L P P A I	P E - - - - -	P P G	I L F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:286)
FGF21C-tail19-53	P G L P P A I	P E - - - - -	P P G	I M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:287)
FGF21C-tail19-54	P G L P P A I	P E - - - - -	P P G	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:288)
FGF21C-tail19-55	P G L P P A I	P E - - - - -	P P G	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:289)
FGF21C-tail19-56	P G L P P A I	P E - - - - -	P E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:290)
FGF21C-tail19-57	P G L P P A I	P E - - - - -	P E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:291)
FGF21C-tail19-58	P G L P P A I	P E - - - - -	P E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:292)
FGF21C-tail19-59	P G L P P A I	P E - - - - -	G H I E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:293)
FGF21C-tail19-60	P G L P P A I	P E - - - - -	G H I E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:294)
FGF21C-tail19-61	P G L P P A I	P E - - - - -	R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:295)
FGF21C-tail19-62	P G L P P A I	P E - - - - -	L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:296)
FGF21C-tail19-63	P G L P P A I	P E B E	D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:297)
FGF21C-tail19-64	P G L P P A I	P E B E	P E D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:298)
FGF21C-tail19-65	P G L V P E E	P E D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:299)	
FGF21C-tail19-66	P G M V P E E	P E D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:300)	
FGF21C-tail19-67	P M V P E E	P E D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:300)	
FGF19C-tail(169)	I P M V P E E	P E D L R G H L E S	D M F S S P L E T D	S M D P F G L V T G	K (SEQ ID NO:300)	

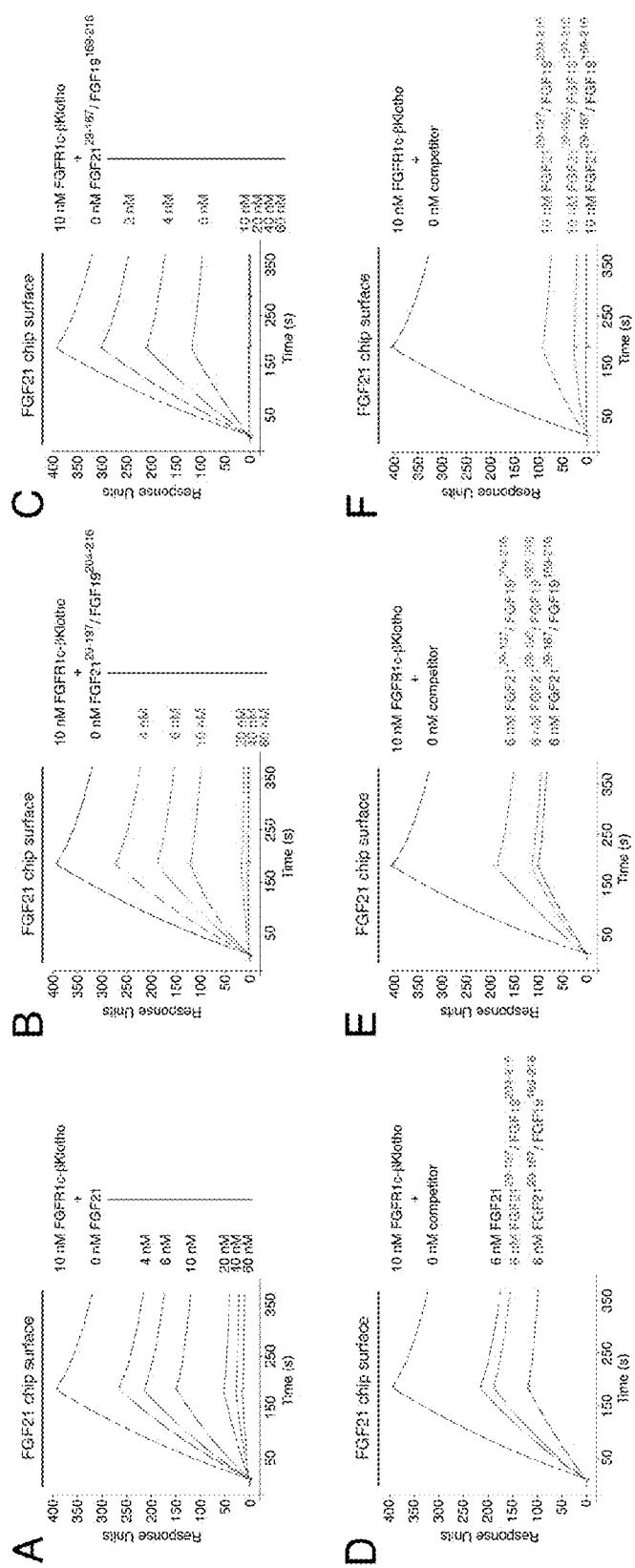
FIG. 12

FGF19C-tail(169) L P M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-1) L P M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-2) L P M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-3) L P M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-4) L P M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-5) L P M V P E E P E D L R G H I P E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-6) L P M V P E E P E D L R - G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-7) L P M V P E E P E D L - G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-8) L P M V P A E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-9) L P M V P P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-10) L P L V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF19C-tail(21-11) L G M V P E E P E D L R G H I L E S D M F S S P L E T D S M D P F G L V T C
 FGF21C-tail(168) P G L P P A L P E - - - - P P G I L A P Q P P D V G S S D P L S M V - G P S Q G R S P S Y A S

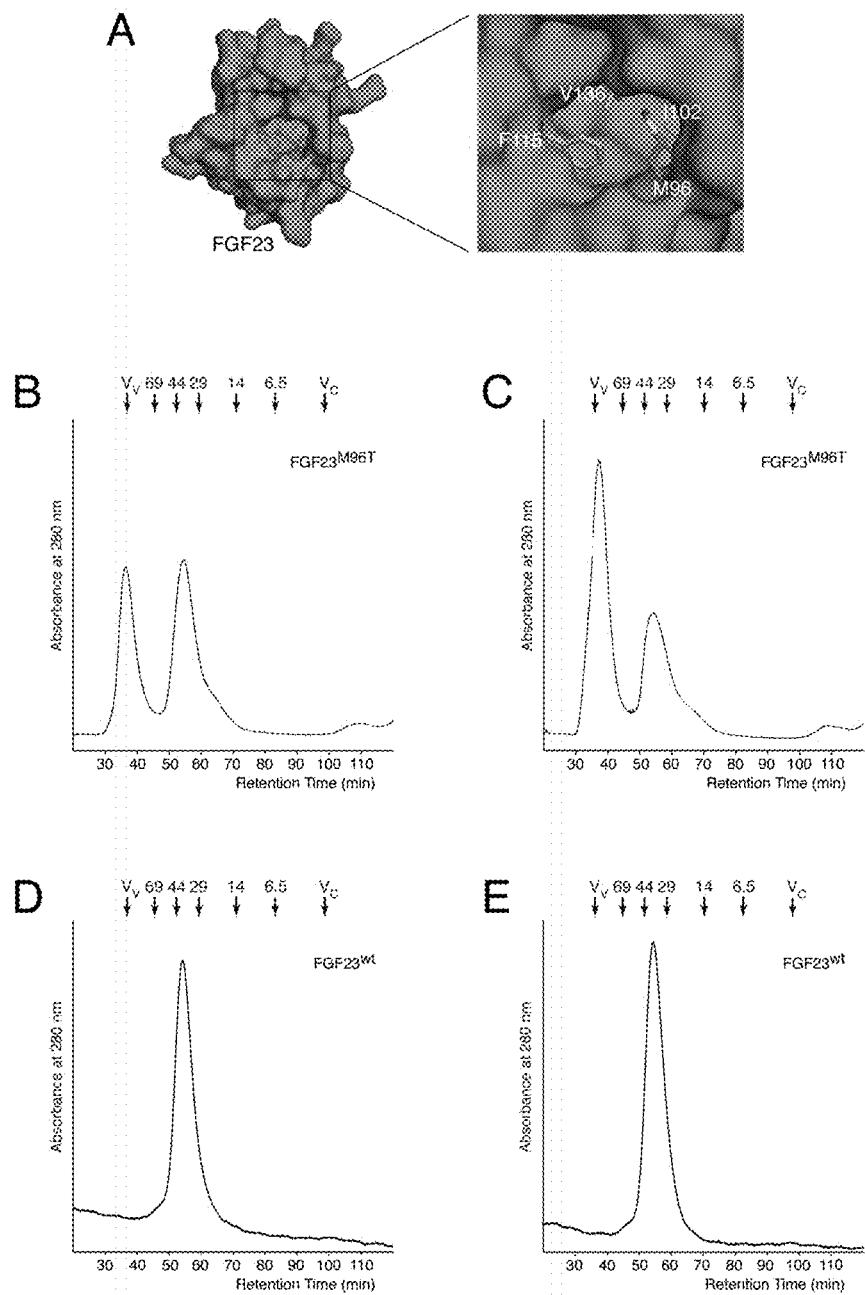
FIG. 13



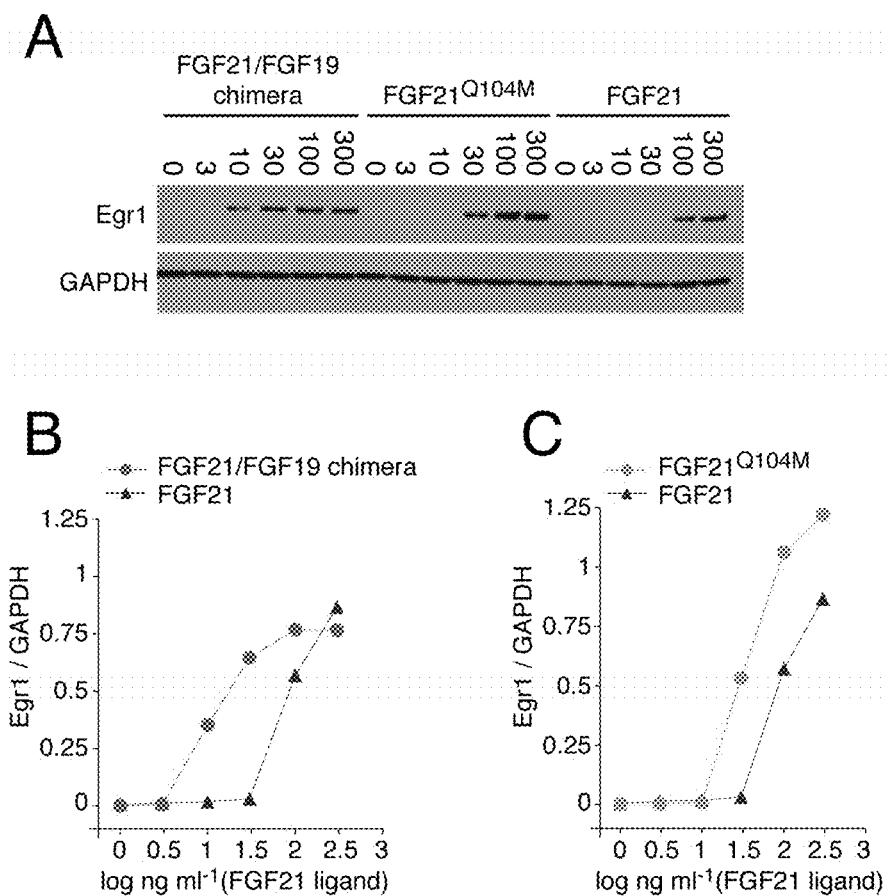
FIGS. 14A-14D



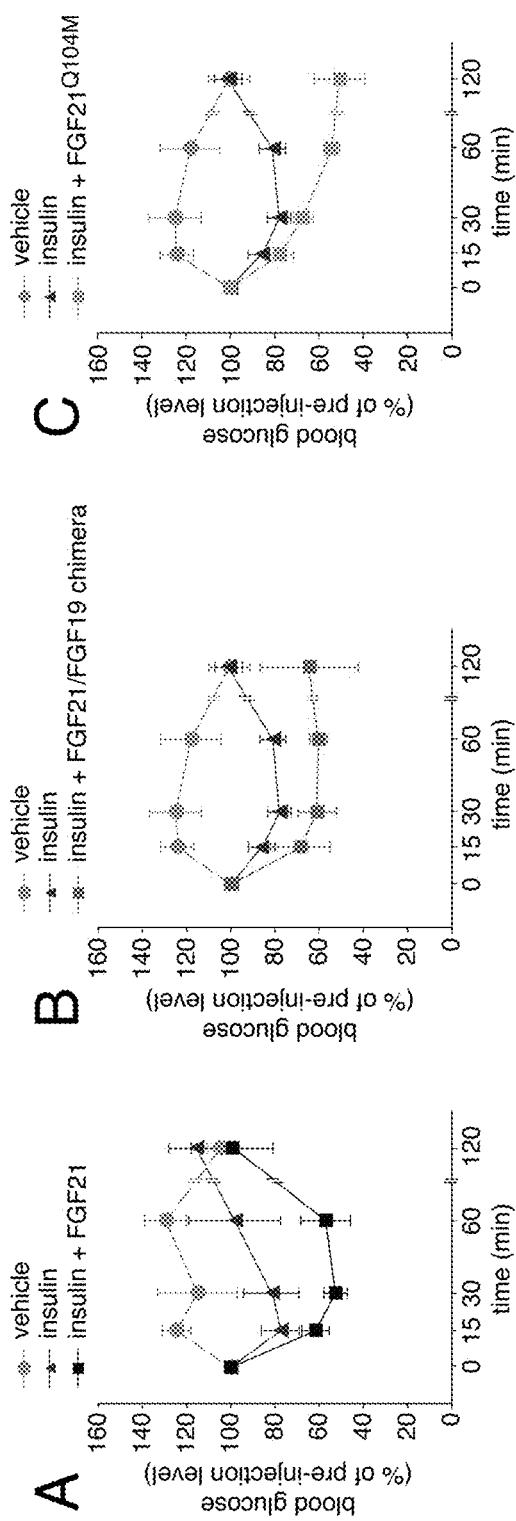
FIGS. 15A-15F



FIGS. 16A-16E



FIGS. 17A-17C

**FIGS. 18A-18C**

1

**CHIMERIC FGF21 PROTEINS WITH
ENHANCED BINDING AFFINITY FOR
 β -KLOTHO FOR THE TREATMENT OF
TYPE II DIABETES, OBESITY, AND
RELATED METABOLIC DISORDERS**

This application claims the benefit of U.S. Provisional Patent Application Ser. No. 61/605,961 filed Mar. 2, 2012, which is hereby incorporated by reference in its entirety.

This invention was made with U.S. government support under DE13686, DK077276, AG019712, DK091392, and DK067158 awarded by the National Institutes of Health. The Government has certain rights in the invention.

FIELD OF THE INVENTION

This invention relates to chimeric FGF21 proteins and their use for the treatment of diabetes, obesity, and related metabolic disorders.

BACKGROUND OF THE INVENTION

Type 2 diabetes is a chronic progressive disorder, which results from end-organ resistance to the action of insulin in combination with insufficient insulin secretion from the pancreas. The metabolic abnormalities associated with insulin resistance and secretory defects, in particular the hyperglycemia, lead over the course of years to extensive irreversible damage to multiple organs including heart, blood vessels, kidney, and eye. Currently, nearly 200 million or 2.9% of the world population have type 2 diabetes (World Health Organization, Diabetes Fact Sheet No 312, January 2011; Wild et al., "Global Prevalence of Diabetes: Estimates for the Year 2000 and Projections for 2030," *Diabetes Care* 27(5):1047-1053 (2004)), and its prevalence is rising at an alarmingly fast pace in parallel with the rise in the prevalence of overweight and obesity (World Health Organization, Obesity and Overweight Fact Sheet No 311, January 2011). Until the end of the 20th century, type 2 diabetes was observed only in adults but what was once known as "adult-onset diabetes" is now also diagnosed in children and adolescents, and this growing incidence can be related to the increase in overweight and obesity among children and adolescents. The prevalence of pre-diabetes, an intermediate metabolic stage between normal glucose homeostasis and diabetes, is even greater than that of type 2 diabetes. Currently, nearly 80 million or 26% of the population in the United States alone have pre-diabetes (Center for Disease Control and Prevention, National Diabetes Fact Sheet 2011), and as such are at high risk for progressing to type 2 diabetes. Type 2 diabetes ranks among the ten leading causes of death worldwide, and the World Health Organization projects that mortality from diabetes (90% of which is type 2) will more than double within the next decade (World Health Organization, Diabetes Fact Sheet No 312, January 2011). Type 2 diabetes also is a major cause of disability. As a consequence of diabetic retinopathy, about 10% of all patients with diabetes in the world develop severe visual impairment and 2% become blind 15 years into the disease (World Health Organization, Diabetes Fact Sheet No 312, January 2011). Diabetic neuropathy, which affects up to half of all patients with diabetes worldwide (World Health Organization, Diabetes Fact Sheet No 312, January 2011), accounts for the majority of nontraumatic lower-limb amputations. Indeed, in its recently published first worldwide report on non-infectious diseases, the World Health Organization considers diabetes, together with other chronic

2

non-infectious diseases like cancer and heart disease, a global economic and social burden, which exceeds that imposed by infectious diseases such as HIV/AIDS.

The current drug therapy for type 2 diabetes is focused on correcting the hyperglycemia in the patients. Although a number of small molecules and biologics with different mechanisms of anti-hyperglycemic action are available for use as mono-therapy or combination therapy, most, if not all of these have limited efficacy, limited tolerability, and significant adverse effects (Moller, "New Drug Targets for Type 2 Diabetes and the Metabolic Syndrome," *Nature* 414 (6865):821-827 (2001)). For example, treatment with sulfonylureas, glinides, thiazolidinediones, or insulin has been associated with weight gain, which is an undesired effect since overweight is considered a driving force in the pathogenesis of type 2 diabetes. Some of these treatments have also been associated with increased risk of hypoglycemia. A limitation specific to the thiazolidinediones is the potential for adverse cardiovascular effects (DeSouza et al., "Therapeutic Targets to Reduce Cardiovascular Disease in Type 2 Diabetes," *Nat Rev Drug Discov* 8(5):361-367 (2009)). A meta-analysis of clinical data on the thiazolidinedione rosiglitazone (Avandia®, which was widely used for the treatment of type 2 diabetes, found that the drug increased the risk of myocardial infarction in patients with type 2 diabetes (Nissen et al., "Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes," *N Engl J Med* 356(24):2457-2471 (2007)). Of all diabetic complications, cardiovascular disease is the main cause of morbidity and mortality in patients with diabetes (World Health Organization, Diabetes Fact Sheet No 312, January 2011; Center for Disease Control and Prevention, National Diabetes Fact Sheet 2011), and hence an aggravation of cardiovascular risk by drug treatment is absolutely unacceptable. In the wake of the debate about the cardiovascular safety of thiazolidinediones, the FDA issued a guidance on evaluating cardiovascular risk in new anti-diabetic therapies to treat type 2 diabetes (Opas A, "Diabetes Drugs Pass Cardiovascular Risk Check," *Nat Rev Drug Discov* 8(5):343-344 (2009)). Meanwhile, thiazolidinediones lost their popularity. Even for glucagon-like peptide-1 agonists, one of the latest class of drugs introduced for the treatment of type 2 diabetes, concerns about safety have been raised, namely the potential for carcinogenicity (Opas A, "Diabetes Drugs Pass Cardiovascular Risk Check," *Nat Rev Drug Discov* 8(5):343-344 (2009)). Therefore, novel therapies that are more effective and safer than existing drugs are needed. Since the currently available drugs do not directly target complications of advanced diabetic disease, especially cardiovascular disease, therapies that are not only effective in lowering blood glucose but also reduce cardiovascular risk factors such as dyslipidemia are particularly desired.

A search conducted by Eli Lilly & Co. for potential novel biotherapeutics to treat type 2 diabetes led to the discovery of fibroblast growth factor (FGF) 21 as a protein that stimulates glucose uptake into adipocytes in an insulin-independent fashion (Kharitonov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6):1627-1635 (2005)). FGF21 has since emerged as a key endocrine regulator not only of glucose metabolism but also of lipid metabolism, and has become one of the most promising drug candidates for the treatment of type 2 diabetes, obesity, and metabolic syndrome. In mouse models of diabetes and obesity, pharmacologic doses of FGF21 lower plasma glucose and increase insulin sensitivity (Kharitonov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest*

115(6):1627-1635 (2005); Coskun et al., "Fibroblast growth factor 21 corrects obesity in mice," *Endocrinology* 149(12): 6018-6027 (2008)). Concurrently, FGF21 lowers plasma triglyceride and cholesterol, enhances lipolysis and suppresses lipogenesis, and accelerates energy expenditure (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6):1627-1635 (2005); Coskun et al., "Fibroblast growth factor 21 corrects obesity in mice," *Endocrinology* 149(12):6018-6027 (2008)). In obese mice, FGF21 causes weight loss, in lean mice, it is weight neutral (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6):1627-1635 (2005); Coskun et al., "Fibroblast growth factor 21 corrects obesity in mice," *Endocrinology* 149(12):6018-6027 (2008)). Thus, FGF21 has some of the most desired characteristics of a drug for the treatment of type 2 diabetes; not only does it improve glycemic control, but also directly affects cardiovascular risk factors, such as hypertriglyceridemia, and reduces obesity, which is considered the single most important promoter of type 2 diabetes. Importantly, FGF21 does not induce hypoglycemia (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6):1627-1635 (2005)), a side effect that can occur with several of the current anti-diabetic therapies, including insulin. Moreover, FGF21 does not exhibit any mitogenic activity in mice (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6):1627-1635 (2005)), ruling out the possibility of a carcinogenic risk. The findings on FGF21 therapy in mouse models of diabetes have been reproduced in diabetic rhesus monkeys (Kharitonenkov et al., "The Metabolic State of Diabetic Monkeys is Regulated by Fibroblast Growth Factor-21," *Endocrinology* 148(2):774-781 (2007)), and are currently followed up with clinical trials in humans (Kharitonenkov et al., "FGF21 Reloaded: Challenges of a Rapidly Growing Field," *Trends Endocrinol Metab* 22(3):81-86 (2011)). However, there is a need for more effective FGF21 therapeutics.

The present invention overcomes these and other deficiencies in the art.

SUMMARY OF THE INVENTION

One aspect of the present invention relates to a chimeric protein that includes an N-terminus coupled to a C-terminus. The N-terminus includes an N-terminal portion of fibroblast growth factor 21 ("FGF21") having a core domain and the C-terminus includes a C-terminal portion of fibroblast growth factor 19 ("FGF19"), where either (i) the N-terminal portion of FGF21 comprises at least one amino acid residue substitution to increase stability of the FGF21 core domain compared to the wild type FGF21; (ii) the C-terminal portion of FGF19 begins at a residue corresponding to any one of residues 169 to 204 of SEQ ID NO:1 and comprises amino acid residues TGLEAV(R/N)SPSFEK (SEQ ID NO: 49); or (iii) both (i) and (ii).

Another aspect of the present invention relates to a pharmaceutical composition that includes a chimeric protein according to the present invention and a pharmaceutically acceptable carrier.

Another aspect of the present invention relates to a method of treating a subject suffering from diabetes, obesity, or metabolic syndrome. This method includes selecting a subject suffering from diabetes, obesity, or metabolic syndrome and administering to this selected subject a therapeutically effective amount of a chimeric protein according to the present invention.

Another aspect of the present invention relates to a method of treating a subject in need of increased FGF21- β Klotho-FGF receptor ("FGFR") complex formation. This method includes selecting a subject in need of increased FGF21- β Klotho-FGFR complex formation and administering to the selected subject a chimeric FGF21 protein, where the chimeric FGF21 protein comprises an FGF21 core domain and a C-terminal portion of FGF19, thereby treating a subject in need of increased FGF21- β Klotho-FGFR complex formation.

Yet another aspect of the present invention relates to a method of causing increased FGF21 receptor agonist- β Klotho-FGFR complex formation. This method comprises providing a cell comprising β Klotho and an FGFR and providing an FGF21 receptor agonist, where the agonist comprises a chimeric protein comprising a C-terminal portion of FGF19. This method also includes contacting the cell and the FGF21 receptor agonist under conditions effective to cause increased FGF21 receptor agonist- β Klotho-FGFR complex formation relative to contacting the cell with FGF21 alone, where the FGF21 has a core domain.

A further aspect of the present invention relates to a method of screening for compounds with enhanced binding affinity for β Klotho suitable for fusion to the C-terminus of an N-terminal portion of FGF21 to generate an FGF21 agonist. The method includes providing FGF21, providing β Klotho, and providing one or more candidate compounds; combining the FGF21, the β Klotho, and the candidate compounds under conditions effective for FGF21 and β Klotho to form a binary complex if present by themselves; and identifying the candidate compounds which diminish binary complex formation, compared to when the candidate compound is absent, as being potentially suitable for fusion to the C-terminus of an N-terminal portion of FGF21 to generate an FGF21 agonist.

Yet a further aspect of the present invention relates to a method of screening for compounds with enhanced binding affinity for the β Klotho-FGFR complex suitable for treatment of diabetes, obesity, or related metabolic disorders. This method includes providing FGF21, providing a binary β Klotho-FGFR complex, and providing one or more candidate compounds. This method also includes combining the FGF21, the binary β Klotho-FGFR complex, and the candidate compounds under conditions effective for the FGF21 and the β Klotho-FGFR complex to form a ternary complex if present by themselves and identifying the candidate compounds which diminish ternary complex formation compared to when the candidate compound is absent as being potentially suitable for treatment of diabetes, obesity, or related metabolic disorders.

FGF21 depends on the co-receptor β Klotho to activate its cognate FGFR (FGFR1c) in its target tissues including white adipose tissue (Ogawa et al., " β Klotho is Required for Metabolic Activity of Fibroblast Growth Factor 21," *Proc Natl Acad Sci USA* 104(18):7432-7437 (2007); Ding et al., " β Klotho is Required for Fibroblast Growth Factor 21 Effects on Growth and Metabolism," *Cell Metab* 16:387-393 (2012), which are hereby incorporated by reference in their entirety). In the course of deciphering the molecular details of how FGF21 forms a signaling complex on the cell surface with FGFR1c and β Klotho, two discoveries were made that provided the basis for the rational design of an FGF21 agonist. It was found that β Klotho promotes binding of FGF21 to its cognate FGFR by engaging ligand and receptor simultaneously through two distinct binding sites (Goetz et al., "Klotho Coreceptors Inhibit Signaling by Paracrine Fibroblast Growth Factor 8 Subfamily Ligands," *Mol Cell*

Biol 32:1944-1954 (2012), which is hereby incorporated by reference in its entirety). β Klotho plays the same role in promoting binding of FGF19, an endocrine regulator of bile acid homeostasis, to its cognate FGFR (Goetz et al., "Klotho Coreceptors Inhibit Signaling by Paracrine Fibroblast Growth Factor 8 Subfamily Ligands," *Mol Cell Biol* 32:1944-1954 (2012), which is hereby incorporated by reference in its entirety). The binding site for β Klotho was mapped on FGF21 and FGF19 to the C-terminal region of each ligand that follows the β -trefoil core domain (Goetz et al., "Klotho Coreceptors Inhibit Signaling by Paracrine Fibroblast Growth Factor 8 Subfamily Ligands," *Mol Cell Biol* 32:1944-1954 (2012), which is hereby incorporated by reference in its entirety). In the course of these studies, it was found that the C-terminal tail peptides of FGF21 and FGF19 share a common binding site on β Klotho, and that the C-terminal tail of FGF19 binds tighter than the C-terminal tail of FGF21 to this site (Goetz et al., "Klotho Coreceptors Inhibit Signaling by Paracrine Fibroblast Growth Factor 8 Subfamily Ligands," *Mol Cell Biol* 32:1944-1954 (2012), which is hereby incorporated by reference in its entirety). As described herein, chimeric FGF21 proteins were made in which C-terminal sequences in FGF21 were replaced with the corresponding sequences of FGF19, which was found to confer greater binding affinity of β Klotho to the chimeras, and, hence, enhance agonistic properties.

In another approach of engineering an FGF21 agonist, residues in the β -trefoil core domain of FGF21 were mutated in order to increase the stability of FGF21. Based on extensive knowledge of the structures of FGF ligands, including the structures of FGF19 and FGF23, Q104 of FGF21 was selected for mutagenesis. As described herein, it was found that replacing Q104 with methionine, which is found in all other FGF ligands at the corresponding position (Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine & Growth Factor Rev* 16(2):107-137 (2005), which is hereby incorporated by reference in its entirety) increases the stability of FGF21 without affecting ligand-binding affinity for receptor. This enhanced affinity for β Klotho, together with the enhanced stability, make these chimeric proteins particularly suitable for use as a therapeutic.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-1D show that the ternary complex of FGF21 with its cognate FGFR and β Klotho coreceptor can be reconstituted in solution using the ectodomains of β Klotho and FGFR1c. FIG. 1A shows a size-exclusion chromatogram of the 1:1 FGFR1c- β Klotho complex. Arrows indicate the retention times of molecular size standards, the void volume (V_v) and the column volume (V_c). Proteins of column peak fractions were resolved on 14% SDS-polyacrylamide gels and stained with Coomassie Blue. FIG. 1B shows a size-exclusion chromatogram of the ternary FGF21-FGFR1c- β Klotho complex. Arrows indicate the retention times of molecular size standards, the void volume (V_v) and the column volume (V_c). Proteins of column peak fractions were resolved on 14% SDS-polyacrylamide gels and stained with Coomassie Blue. FIG. 1C shows a representative surface plasmon resonance (SPR) sensorgram illustrating binding of FGF21 to the binary FGFR1c- β Klotho complex. FGF21 was immobilized on a biosensor chip, and increasing concentrations of FGFR1c- β Klotho complex were passed over the chip. FIG. 1D shows a representative SPR sensorgram illustrating no interaction between FGF21 and the binary FGFR1c- α Klotho complex. FGF21 was immobilized

on a biosensor chip, and two concentrations of FGFR1c- α Klotho complex were passed over the chip.

FIGS. 2A-2B show that the ternary complex of FGF19 with its cognate FGFR and β Klotho coreceptor can be reconstituted in solution using the ectodomains of β Klotho and FGFR4. FIG. 2A shows a size-exclusion chromatogram of the 1:1 FGFR4- β Klotho complex. Arrows indicate the retention times of molecular size standards, the void volume (V_v) and the column volume (V_c). Proteins of column peak fractions were resolved on 14% SDS-polyacrylamide gels and stained with Coomassie Blue. FIG. 2B shows a size-exclusion chromatogram of the ternary FGF19-FGFR4- β Klotho complex. Arrows indicate the retention times of molecular size standards, the void volume (V_v) and the column volume (V_c). Proteins of column peak fractions were resolved on 14% SDS-polyacrylamide gels and stained with Coomassie Blue.

FIGS. 3A-3G show the FGFR binding specificity profile of β Klotho. FIG. 3A shows an overlay of SPR sensorgrams of FGFR1c binding to β Klotho, and fitted saturation binding curve. FIG. 3B shows an overlay of SPR sensorgrams of FGFR2c binding to β Klotho, and fitted saturation binding curve. FIG. 3C shows an overlay of SPR sensorgrams of FGFR3c binding to β Klotho. FIG. 3D shows an overlay of SPR sensorgrams of FGFR4 binding to β Klotho, and fitted saturation binding curve. FIG. 3E shows an overlay of SPR sensorgrams of FGFR1b binding to β Klotho. FIG. 3F shows an overlay of SPR sensorgrams of FGFR2b binding to β Klotho. FIG. 3G shows an overlay of SPR sensorgrams of FGFR3b binding to β Klotho. β Klotho ectodomain was immobilized on a biosensor chip, and increasing concentrations of the ligand-binding domain of each of the seven principal human FGFRs were passed over the chip. Where binding was observed, the dissociation constant (K_D) was calculated from the saturation binding curve. The data shown in FIGS. 3A-G are representative of two to five independent experiments.

FIGS. 4A-4E show that β Klotho contains a high affinity binding site for FGF19 and FGF21. FIG. 4A shows an overlay of SPR sensorgrams illustrating β Klotho binding to FGF19. FGF19 was immobilized on a biosensor chip, and increasing concentrations of β Klotho ectodomain were passed over the chip. FIG. 4B shows an overlay of SPR sensorgrams illustrating β Klotho binding to FGF21. FGF21 was immobilized on a biosensor chip, and increasing concentrations of β Klotho ectodomain were passed over the chip. Note that for any given concentration of β Klotho, the binding response is greater on the FGF19 chip surface than on the FGF21 chip surface. Also note that the FGF19- β Klotho complex dissociates more slowly than the FGF21- β Klotho complex (compare the dissociation phases of the sensorgrams shown in (A) and (B)). FIG. 4C shows an overlay of SPR sensorgrams illustrating no interaction between β Klotho and FGF23. FGF23 was immobilized on a biosensor chip, and increasing concentrations of β Klotho ectodomain were passed over the chip. FIG. 4D shows an overlay of SPR sensorgrams illustrating no interaction between α Klotho and FGF19. FGF19 was immobilized on a biosensor chip, and increasing concentrations of α Klotho ectodomain were passed over the chip. FIG. 4E shows an overlay of SPR sensorgrams illustrating no interaction between α Klotho and FGF21. FGF21 was immobilized on a biosensor chip, and increasing concentrations of α Klotho ectodomain were passed over the chip. The data shown in FIGS. 4A-E are representative of two to three independent experiments.

FIGS. 5A-5G show that the C-terminal tail peptides of FGF19 and FGF21 share a common binding site on β Klotho, and that the C-terminal tail peptide of FGF19 has greater affinity for this site than the C-terminal tail peptide of FGF21. FIG. 5A shows an alignment of the C-terminal tail sequences of human FGF19 (SEQ ID NO:1), FGF21 (SEQ ID NO:100), and FGF23 (SEQ ID NO:223). Residue numbers are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. Residues that are identical between FGF19 and FGF21 are shaded gray. Note that 40% of these residues map to the most C-terminal sequence. FIG. 5B shows an overlay of SPR sensograms illustrating inhibition by the FGF19 C-terminal tail peptide (M171 to K216 of SEQ ID NO:1; FGF19^{C-tail}) of β Klotho binding to FGF19. FGF19 was immobilized on a biosensor chip, and mixtures of a fixed concentration of β Klotho ectodomain with increasing concentrations of FGF19^{C-tail} were passed over the chip. FIG. 5C shows an overlay of SPR sensograms illustrating inhibition by the FGF21 C-terminal tail peptide (P168 to S209 of SEQ ID NO:100; FGF21^{C-tail}) of β Klotho binding to FGF19. Mixtures of a fixed concentration of β Klotho ectodomain with increasing concentrations of FGF21^{C-tail} were passed over a biosensor chip onto which FGF19 had been immobilized. FIG. 5D shows an overlay of SPR sensograms illustrating no inhibition by the FGF23 C-terminal tail peptide (S180 to I251 of SEQ ID NO:223; FGF23^{C-tail}) of β Klotho binding to FGF19. β Klotho ectodomain and FGF23^{C-tail} were mixed at a molar ratio of 1:2, and the mixture was injected over a biosensor chip onto which FGF19 had been immobilized. FIG. 5E shows an overlay of SPR sensograms illustrating inhibition by the FGF21 C-terminal tail peptide (P168 to S209 of SEQ ID NO:100; FGF21^{C-tail}) of β Klotho binding to FGF21. FGF21 was immobilized on a biosensor chip, and mixtures of a fixed concentration of β Klotho ectodomain with increasing concentrations of FGF21^{C-tail} were passed over the chip. FIG. 5F shows an overlay of SPR sensograms illustrating inhibition by the FGF19 C-terminal tail peptide (M171 to K216 of SEQ ID NO:1; FGF19^{C-tail}) of β Klotho binding to FGF21. Mixtures of a fixed concentration of β Klotho ectodomain with increasing concentrations of FGF19^{C-tail} were passed over a biosensor chip onto which FGF21 had been immobilized. FIG. 5G shows an overlay of SPR sensograms illustrating no inhibition by the FGF23 C-terminal tail peptide (S180 to I251 of SEQ ID NO:223; FGF23^{C-tail}) of β Klotho binding to FGF21. β Klotho ectodomain and FGF23^{C-tail} were mixed at a molar ratio of 1:2, and the mixture was injected over a biosensor chip onto which FGF21 had been immobilized. The data shown in FIGS. 5B-G are representative of two to three independent experiments.

FIGS. 6A-6B show that the C-terminal tail peptides of FGF19 and FGF21 are interchangeable in inhibiting the signaling of FGF19. FIG. 6A shows an immunoblot analysis for phosphorylation of FRS2 α (pFRS2 α) and 44/42 MAP kinase (p44/42 MAPK) in the rat hepatoma cell line H4IIE, which had been stimulated with either FGF19 or FGF19^{C-tail} alone, or with mixtures of FGF19 with increasing concentrations of FGF19^{C-tail}. Numbers above the lanes give the amounts of protein/peptide added in ng ml⁻¹. To control for equal sample loading, the protein blots were probed with an antibody recognizing both phosphorylated and nonphosphorylated (total) 44/42 MAP kinase (44/42 MAPK). FIG. 6B shows an immunoblot analysis for phosphorylation of FRS2 α (pFRS2 α) and 44/42 MAP kinase (p44/42 MAPK) in the rat hepatoma cell line H4IIE, which had been stimu-

lated with either FGF19 or FGF21^{C-tail} alone, or with mixtures of FGF19 with increasing concentrations of FGF21^{C-tail}. Numbers above the lanes give the amounts of protein/peptide added in ng ml⁻¹. To control for equal sample loading, the protein blots were probed with an antibody recognizing both phosphorylated and nonphosphorylated (total) 44/42 MAP kinase (44/42 MAPK). The data shown in FIGS. 6A-B are representative of two independent experiments. Note that while FGF21^{C-tail} can inhibit FGF19 signaling in H4IIE cells, this cell line is otherwise not responsive to FGF21.

FIGS. 7A-7C show that a FGF21/FGF19 chimera has enhanced binding affinity for β Klotho. FIG. 7A shows an overlay of SPR sensograms illustrating inhibition by FGF21 in solution of β Klotho binding to FGF21 immobilized on a biosensor chip. Increasing concentrations of FGF21 were mixed with a fixed concentration of β Klotho ectodomain, and the mixtures were passed over a FGF21 chip. FIG. 7B shows an overlay of SPR sensograms illustrating inhibition by the FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera of β Klotho binding to FGF21 immobilized on a biosensor chip. Increasing concentrations of FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera were mixed with a fixed concentration of β Klotho ectodomain, and the mixtures were passed over a FGF21 chip. FIG. 7C shows an overlay of SPR sensograms illustrating inhibition by the FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera or FGF21 of β Klotho binding to immobilized FGF21. The figure was created from the data shown in FIGS. 7A-B, which are representative of two independent experiments.

FIGS. 8A-8B show a schematic of one of the FGF21/FGF19 chimeras claimed in this invention and an alignment of the primary sequences of FGF19 and FGF21. FIG. 8A shows a schematic of the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera claimed as an FGF21 agonist herein. The amino acid boundaries of each component of the chimera are labeled. The FGF19 portion of the chimera is shaded gray. FIG. 8B shows a sequence alignment of human FGF19 and FGF21. Residue numbers are in parenthesis to the left of the alignment. The secondary structure elements known for FGF19 (β 1- β 12, α 11) are indicated above the alignment, and FGF19 residues containing these elements are boxed. A dashed line across the alignment marks the junction between the FGF homology core domain and the C-terminal tail of FGF19 and FGF21. Gaps (dashes) were introduced to optimize the sequence alignment. Residues that are identical between FGF19 and FGF21 are shaded gray.

FIG. 9 shows a sequence alignment of the FGF homology core domain and its N-terminal extension of FGF21 orthologs (including human (SEQ ID NO: 100), orangutan (SEQ ID NO: 101), chimpanzee (SEQ ID NO: 102), elephant (SEQ ID NO: 115), dog (SEQ ID NO: 103), pig (SEQ ID NO: 116), bovine (SEQ ID NO: 104), horse (SEQ ID NO: 105), panda (SEQ ID NO: 106), rabbit (SEQ ID NO: 107), squirrel (SEQ ID NO: 114), rat (SEQ ID NO: 119), and mouse (SEQ ID NO: 120)). Residue numbers are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. Ortholog residues identical to human FGF21 (SEQ ID NO: 100) are shaded gray. This illustrates the high degree of sequence conservation among mammals.

FIG. 10 shows a sequence alignment of the C-terminal tail of FGF19 orthologs (including human (SEQ ID NO: 1), gorilla (SEQ ID NO: 2), chimpanzee (SEQ ID NO: 3), gibbon (SEQ ID NO: 6), rhesus monkey (SEQ ID NO: 4), orangutan (SEQ ID NO: 5), marmoset (SEQ ID NO: 7), mouse lemur (SEQ ID NO: 8), sloth (SEQ ID NO: 9), panda

(SEQ ID NO: 10), pig (SEQ ID NO: 11), bovine (SEQ ID NO: 12), dog (SEQ ID NO: 13), rabbit (SEQ ID NO: 14), megabat (SEQ ID NO: 15), dolphin (SEQ ID NO: 16), microbat (SEQ ID NO: 17), platypus (SEQ ID NO: 18), opossum (SEQ ID NO: 19), anole lizard (SEQ ID NO: 20), pika (SEQ ID NO: 21), guinea pig (SEQ ID NO: 22), tree shrew (SEQ ID NO: 23), rat (SEQ ID NO: 24), mouse (SEQ ID NO: 25), chicken (SEQ ID NO: 26), zebra finch (SEQ ID NO: 27), zebrafish (SEQ ID NO: 28), and frog (SEQ ID NO: 29)). Residue numbers are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. Ortholog residues identical to human FGF19 are shaded gray.

FIG. 11 shows an alignment of the C-terminal tail sequences of human FGF21 (SEQ ID NO: 100), FGF19 (SEQ ID NO: 1), and variants of FGF21 harboring a single amino acid substitution or insertion for a residue unique to FGF19. Residue numbers for the sequences of native or wildtype FGF21 (SEQ ID NO: 100) and FGF19 (SEQ ID NO: 1) are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. In the sequence of native or wildtype FGF19 (SEQ ID NO: 1), residues unique to FGF19 are bold and boxed, and in the sequences of the variants of the FGF21 C-terminal tail, introduced FGF19 residues are highlighted in the same manner.

FIG. 12 shows an alignment of the C-terminal tail sequences of human FGF21 (SEQ ID NO: 100), FGF19 (SEQ ID NO: 1), and variants of FGF21 in which residues unique to FGF19 progressively replace the corresponding residues of FGF21 or are inserted into the FGF21 sequence. Residue numbers for the sequences of native FGF21 (SEQ ID NO: 100) and FGF19 (SEQ ID NO: 1) are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. In the sequence of native FGF19 (SEQ ID NO: 1), residues unique to FGF19 are bold and boxed, and in the sequences of variants of the FGF21 C-terminal tail, introduced FGF19 residues are highlighted in the same manner.

FIG. 13 shows an alignment of the C-terminal tail sequences of human FGF19 (SEQ ID NO: 1), FGF21 (SEQ ID NO: 100), and variants of FGF19 harboring a single amino acid deletion or substitution for a residue unique to FGF21. Residue numbers for the sequences of native FGF19 (SEQ ID NO: 1) and FGF21 (SEQ ID NO: 100) are in parenthesis to the left of the alignment. Gaps (dashes) were introduced to optimize the alignment. In the sequence of native or wildtype FGF21 (SEQ ID NO: 100), residues unique to FGF21 are bold and boxed, and in the sequences of the variants of the FGF 19 C-terminal tail, introduced FGF21 residues are also bold and boxed and deleted FGF19 residues are indicated by a dash (bold and boxed).

FIGS. 14A-14D show that substitution of the last three residues at the C-terminus of FGF21 for the corresponding residues of FGF 19 reduces the binding affinity of FGF21 for the FGFR1c- β Klotho complex. FIG. 14A shows an overlay of SPR sensorgrams illustrating inhibition by FGF21 in solution of FGFR1c- β Klotho binding to FGF21 immobilized on a biosensor chip. Increasing concentrations of FGF21 were mixed with a fixed concentration of FGFR1c- β Klotho complex, and the mixtures were passed over a FGF21 chip. FIG. 14B shows an overlay of SPR sensorgrams illustrating inhibition by single mutant FGF21 or wild-type FGF21 of FGFR1c- β Klotho binding to FGF21 immobilized on a biosensor chip. FIG. 14C shows an overlay of SPR sensorgrams illustrating inhibition by triple mutant FGF21 or wild-type FGF21 of FGFR1c- β Klotho

binding to FGF21 immobilized on a biosensor chip. FIG. 14D shows an overlay of SPR sensorgrams illustrating inhibition by single mutant FGF21 or triple mutant FGF21 of FGFR1c- β Klotho binding to immobilized FGF21. In the experiments shown in FIGS. 14B-14C, FGFR1c- β Klotho complex was mixed with either mutant FGF21 or wild-type FGF21 at a molar ratio of 1:2 or 1:6, and the mixtures were injected over a FGF21 chip. The data shown in FIGS. 14A-14C are representative of two to three independent experiments. FIG. 14D was created from the data shown in FIGS. 14B-14C. Note that the mutants are less potent than wild-type FGF21 at inhibiting binding of the FGFR1c- β Klotho complex to immobilized FGF21. Also note that the triple mutant exhibits a greater reduction of inhibitory potency than the single mutant.

FIGS. 15A-15F show that FGF21/FGF19 chimeras have enhanced binding affinity for the FGFR1c- β Klotho complex. FIG. 15A shows an overlay of SPR sensorgrams illustrating inhibition by FGF21 in solution of FGFR1c- β Klotho binding to FGF21 immobilized on a biosensor chip. FIG. 15B shows an overlay of SPR sensorgrams illustrating inhibition by the FGF21²⁹⁻⁹⁷/FGF19²⁰⁴⁻²¹⁶ chimera of FGFR1c- β Klotho binding to FGF21 immobilized on a biosensor chip. FIG. 15C shows an overlay of SPR sensorgrams illustrating inhibition by the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera of FGFR1c- β Klotho binding to FGF21 immobilized on a biosensor chip. In the experiments shown in FIGS. 15A-15C, increasing concentrations of either an FGF21/FGF19 chimera or wild-type FGF21 were mixed with a fixed concentration of FGFR1c- β Klotho complex, and the mixtures were passed over a FGF21 chip. FIG. 15D shows an overlay of SPR sensorgrams illustrating inhibition by either of two FGF21/FGF19 chimeras or wild-type FGF21 of FGFR1c- β Klotho binding to immobilized FGF21. FIG. 15E shows an overlay of SPR sensorgrams illustrating inhibition by either of three FGF21/FGF19 chimeras of FGFR1c- β Klotho binding to immobilized FGF21. FIG. 15F shows an overlay of SPR sensorgrams illustrating inhibition by either of three FGF21/FGF19 chimeras of FGFR1c- β Klotho binding to immobilized FGF21. The data shown in FIGS. 15A-15C are representative of two to three independent experiments. FIGS. 15D-15F were created from the data shown in FIGS. 15A-15C. Included in FIGS. 15E-15F are SPR sensorgrams obtained from injecting mixtures of the FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera with the FGFR1c- β Klotho complex over a FGF21 chip.

FIGS. 16A-16E show that substitution of methionine at position 96 for threonine in FGF23 (SEQ ID NO: 224), as it occurs in Familial Tumoral Calcinosis (Chefetz et al., "A Novel Homozygous Missense Mutation in FGF23 Causes Familial Tumoral Calcinosis Associated with Disseminated Visceral Calcification," *Hum Genet.* 118(2):261-266 (2005), which is hereby incorporated by reference in its entirety), destabilizes the FGF23 protein. FIG. 16A shows a molecular surface representation of the FGF23 crystal structure (PDB ID: 2P39; Goetz et al., "Molecular Insights into the Klotho-Dependent, Endocrine Mode of Action of Fibroblast Growth Factor 19 Subfamily Members," *Mol Cell Biol.* 27:3417-3428 (2007), which is hereby incorporated by reference in its entirety). A close-up view into the hydrophobic interior core of FGF23's β -trefoil core domain showing some of the key hydrophobic side chains is shown on the right, and a view of the whole structure is shown on the left. Note that M96 makes numerous hydrophobic contacts with its neighboring residues such as I102, F115, and V136 in the β -trefoil core of FGF23. The M96T substitution would weaken these hydrophobic contacts leading to thermal instability of the

FGF23 protein. FIG. 16B shows a size-exclusion chromatogram of the M96T mutant of FGF23 analyzed immediately after Ni-chelating affinity purification. FIG. 16C shows a size-exclusion chromatogram of the M96T mutant of FGF23 analyzed following incubation at 4° C. for 24 hours. FIG. 16D shows a size-exclusion chromatogram of wild-type FGF23 immediately following protein purification. FIG. 16E shows a size-exclusion chromatogram of purified wild-type FGF23 following incubation at 4° C. for 24 hours. Arrows in FIGS. 16B-16E indicate the retention times of molecular size standards, the void volume (V_0) and the column volume (V_c). Note that in contrast to wild-type FGF23, there is a substantial increase in the portion of M96T mutant protein eluting in the void volume indicating that the mutant protein unfolds over time.

FIGS. 17A-17C show that a FGF21/FGF19 chimera and a single mutant of FGF21 harboring Q104M substitution in the core domain act as FGF21 agonists in a cell-based assay. FIG. 17A shows an immunoblot analysis for early growth response 1 (Egr1) expression in HEK293-βKlotho cells stimulated with FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera, FGF21^{Q104M} mutant, or wild-type FGF21. Numbers above the lanes give the amounts of protein added in ng ml⁻¹. To control for equal sample loading, the protein blots were probed with an antibody to glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The data are representative of two independent experiments. FIG. 17B shows the dose-response curve for induction of Egr1 protein expression in HEK293-βKlotho cells by the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera or wild-type FGF21. The intensity of the protein bands on the immunoblots shown in FIG. 17A was quantified and the ratio of Egr1 to GAPDH was calculated. The ratio of Egr1 to GAPDH is plotted as a function of FGF21 ligand concentration. FIG. 17C shows the dose-response curve for induction of Egr1 protein expression in HEK293-βKlotho cells by the FGF21^{Q104M} mutant or wild-type FGF21. The intensity of the protein bands on the immunoblots shown in FIG. 17A was quantified and the ratio of Egr1 to GAPDH was calculated. The ratio of Egr1 to GAPDH is plotted as a function of FGF21 ligand concentration.

FIGS. 18A-18C show that a FGF21/FGF19 chimera and a single mutant of FGF21 harboring Q104M substitution in the core domain exhibit prolonged potentiating effects on insulin-induced hypoglycemia. FIG. 18A shows changes in blood glucose levels in healthy mice in response to injection of insulin alone or insulin plus FGF21 or vehicle. FIG. 18B shows changes in blood glucose levels in healthy mice in response to injection of insulin alone or insulin plus FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera or vehicle. FIG. 18C shows changes in blood glucose levels in healthy mice in response to injection of insulin alone or insulin plus FGF21^{Q104M} mutant or vehicle. Blood glucose concentrations were measured before and at the indicated time points after the injection of protein(s) or vehicle. Blood glucose concentrations are expressed as percent of pre-injection values. Error bars denote standard deviation from mean.

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention relates to a chimeric protein that includes an N-terminus coupled to a C-terminus. The N-terminus includes an N-terminal portion of fibroblast growth factor 21 (“FGF21”) having a core domain and the C-terminus includes a C-terminal portion of fibroblast growth factor 19 (“FGF19”), where either (i) the N-terminal portion of FGF21 comprises at least one amino acid residue

substitution to increase stability of the FGF21 core domain compared to the wild type FGF21; (ii) the C-terminal portion of FGF19 begins at a residue corresponding to any one of residues 169 to 204 of SEQ ID NO:1 and comprises 5 amino acid residues TGLEAV(R/N)SPSFEK (SEQ ID NO:49); or (iii) both (i) and (ii).

As used herein, the terms “chimeric polypeptide” and “chimeric protein” encompass a polypeptide having a sequence that includes at least a portion of a full-length sequence of first polypeptide sequence and at least a portion of a full-length sequence of a second polypeptide sequence, where the first and second polypeptides are different polypeptides. A chimeric polypeptide also encompasses polypeptides that include two or more non-contiguous portions derived from the same polypeptide. A chimeric polypeptide or protein also encompasses polypeptides having at least one substitution, wherein the chimeric polypeptide includes a first polypeptide sequence in which a portion of the first polypeptide sequence has been substituted by a portion of a 10 second polypeptide sequence.

As used herein, the term “N-terminal portion” of a given polypeptide sequence is a contiguous stretch of amino acids of the given polypeptide sequence that begins at or near the N-terminal residue of the given polypeptide sequence. An 15 N-terminal portion of the given polypeptide can be defined by a contiguous stretch of amino acids (e.g., a number of amino acid residues). Similarly, the term “C-terminal portion” of a given polypeptide sequence is a contiguous length of the given polypeptide sequence that ends at or near the 20 C-terminal residue of the given polypeptide sequence. A C-terminal portion of the given polypeptide can be defined by a contiguous stretch of amino acids (e.g., a number of 25 amino acid residues).

The term “portion,” when used herein with respect to a 30 given polypeptide sequence, refers to a contiguous stretch of amino acids of the given polypeptide’s sequence that is shorter than the given polypeptide’s full-length sequence. A portion of a given polypeptide may be defined by its first 35 position and its final position, in which the first and final 40 positions each correspond to a position in the sequence of the given full-length polypeptide. The sequence position corresponding to the first position is situated N-terminal to the sequence position corresponding to the final position. The sequence of the portion is the contiguous amino acid 45 sequence or stretch of amino acids in the given polypeptide that begins at the sequence position corresponding to the first position and ending at the sequence position corresponding to the final position. A portion may also be defined by reference to a position in the given polypeptide sequence 50 and a length of residues relative to the referenced position, whereby the sequence of the portion is a contiguous amino acid sequence in the given full-length polypeptide that has the defined length and that is located in the given polypeptide in reference to the defined position.

55 As noted above, a chimeric protein according to the present invention may include an N-terminus coupled to a C-terminus. N-terminus and C-terminus are used herein to refer to the N-terminal region or portion and the C-terminal region or portion, respectively, of the chimeric protein of the 60 present invention. In some embodiments of the present invention, the C-terminal portion and the N-terminal portion of the chimeric protein of the present invention are contiguously joined. In alternative embodiments, the C-terminal portion and the N-terminal portion of the chimeric protein of 65 the present invention are coupled by an intervening spacer. In one embodiment, the spacer may be a polypeptide sequence of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid

residues. In some embodiments, the C-terminal portion and/or the N-terminal portion of the chimeric protein of the present invention may include additional portion(s) coupled to the C-terminal residue and/or the N-terminal residue of the chimeric protein of the present invention, respectively. In some embodiments, the additional portion(s) may be a polypeptide sequence of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid residues. In some embodiments, the N-terminal portion and/or the C-terminal portion having such additional portion(s) will maintain the activity of the corresponding naturally occurring N-terminal portion of FGF21 and/or C-terminal portion of FGF19, respectively. In some embodiments, the N-terminal portion and/or the C-terminal portion having such additional portion(s) will have enhanced and/or prolonged activity compared to the corresponding naturally occurring N-terminal portion of FGF21 and/or C-terminal portion of FGF 19, respectively. In other embodiments, the C-terminal portion and/or the N-terminal portion of the chimeric protein of the present invention do not include any additional portion(s) coupled to the C-terminal residue and/or the N-terminal residue of the chimeric protein of the present invention, respectively.

As described by Goetz et al. (Goetz et al., "Molecular Insights into the Klotho-Dependent, Endocrine Mode of Action of Fibroblast Growth Factor 19 Subfamily Members," *Mol Cell Biol* 34:17-3428 (2007), which is hereby incorporated by reference in its entirety), the mammalian fibroblast growth factor (FGF) family comprises 18 polypeptides (FGF1 to FGF10 and FGF16 to FGF23), which participate in a myriad of biological processes during embryo genesis, including but not limited to gastrulation, body plan formation, somitogenesis, and morphogenesis of essentially every tissue/organ such as limb, lung, brain, and kidney (Bottcher et al., "Fibroblast Growth Factor Signaling During Early Vertebrate Development," *Endocr Rev* 26:63-77 (2005), and Thisse et al., "Functions and Regulations of Fibroblast Growth Factor Signaling During Embryonic Development," *Dev Biol* 287:390-402 (2005), which are hereby incorporated by reference in their entirety).

FGFs execute their biological actions by binding to, dimerizing, and activating FGFR tyrosine kinases, which are encoded by four distinct genes (Fgfr1 to Fgfr4). Prototypical FGFRs consist of an extracellular domain composed of three immunoglobulin-like domains, a single-pass transmembrane domain, and an intracellular domain responsible for the tyrosine kinase activity (Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine Growth Factor Rev* 16:107-137 (2005), which is hereby incorporated by reference in its entirety).

The number of principal FGFRs is increased from four to seven due to a major tissue-specific alternative splicing event in the second half of the immunoglobulin-like domain 3 of FGFR1 to FGFR3, which creates epithelial lineage-specific "b" and mesenchymal lineage-specific "c" isoforms (Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine Growth Factor Rev* 16:107-137 (2005) and Ornitz et al., "Fibroblast Growth Factors," *Genome Biol* 2(3):reviews3005.1-reviews3005.12 (2001), which are hereby incorporated by reference in their entirety). Generally, the receptor-binding specificity of FGFs is divided along this major alternative splicing of receptors whereby FGFRb-interacting FGFs are produced by epithelial cells and FGFRc-interacting FGFs are produced by mesenchymal cells (Ornitz et al., "Fibroblast Growth Factors," *Genome Biol* 2(3):reviews3005.1-reviews3005.12 (2001), which is hereby incorporated by reference in its entirety). These reciprocal expression patterns of FGFs and

FGFRs result in the establishment of specific paracrine FGF signaling loops between the epithelium and the mesenchyme, which is essential for proper organogenesis and patterning during embryonic development as well as tissue homeostasis in the adult organism.

Based on sequence homology and phylogenetic and structural considerations, the eighteen mammalian FGFs are grouped into six subfamilies (Itoh et al., "Fibroblast growth factors: from molecular evolution to roles in development, metabolism, and disease," *J Biochem* 149:121-130 (2011); Mohammadi et al., "Structural basis for fibroblast growth factor receptor activation," *Cytokine Growth Factor Rev* 16:107-137 (2005), which are hereby incorporated by reference in its entirety). The FGF core homology domain (approximately 120 amino acids long) is flanked by N- and C-terminal sequences that are highly variable in both length and primary sequence, particularly among different FGF subfamilies. The core region of FGF19 shares the highest sequence identity with FGF21 (38%) and FGF23 (36%), and therefore, these ligands are considered to form a subfamily.

Based on mode of action, the eighteen mammalian FGFs are grouped into paracrine-acting ligands (five FGF subfamilies) and endocrine-acting ligands (one FGF subfamily) comprising FGF19, FGF21 and FGF23 (Itoh and Ornitz, "Fibroblast Growth Factors: From Molecular Evolution to Roles in Development, Metabolism and Disease," *J. Biochem.* 149:121-130 (2011); Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine Growth Factor Rev.* 16:107-137 (2005), which are hereby incorporated by reference in their entirety).

Paracrine FGFs direct multiple processes during embryogenesis, including gastrulation, somitogenesis, organogenesis, and tissue patterning (Itoh and Ornitz, "Fibroblast Growth Factors: From Molecular Evolution to Roles in Development, Metabolism and Disease," *J. Biochem.* 149:121-130 (2011); Bottcher and Niehrs, "Fibroblast Growth Factor Signaling During Early Vertebrate Development," *Endocr. Rev.* 26:63-77 (2005); Thisse et al., "Functions and Regulations of Fibroblast Growth Factor Signaling During Embryonic Development," *Dev. Biol.* 287:390-402 (2005), which are hereby incorporated by reference in their entirety), and also regulate tissue homeostasis in the adult (Hart et al., "Attenuation of FGF Signalling in Mouse Beta-cells Leads to Diabetes," *Nature* 408:864-868 (2000); Jonker et al., "A PPAR γ -FGF1 Axis is Required for Adaptive Adipose Remodelling and Metabolic Homeostasis," *Nature* 485:391-394 (2012), which is hereby incorporated by reference in its entirety).

Endocrine FGFs control major metabolic processes such as bile acid homeostasis (Inagaki et al., "Fibroblast Growth Factor 15 Functions as an Enterohepatic Signal to Regulate Bile Acid Homeostasis," *Cell Metab.* 2:217-225 (2005), which is hereby incorporated by reference in its entirety), and hepatic glucose and protein metabolism (Kir et al., "FGF19 as a Postprandial, Insulin-Independent Activator of Hepatic Protein and Glycogen Synthesis," *Science* 331:1621-1624 (2011); Potthoff et al., "FGF15/19 Regulates Hepatic Glucose Metabolism by Inhibiting the CREB-PGC-1 α Pathway," *Cell Metab.* 13:729-738 (2011), which are hereby incorporated by reference in their entirety) (FGF19), glucose and lipid metabolism (Badman et al., "Hepatic Fibroblast Growth Factor 21 Is Regulated by PPAR α and Is a Key Mediator of Hepatic Lipid Metabolism in Ketotic States," *Cell Metab.* 5:426-437 (2007); Inagaki et al., "Endocrine Regulation of the Fasting Response by PPAR α -mediated Induction of Fibroblast Growth Factor 21," *Cell Metab.* 5:415-425 (2007); Kharitonov et al., "FGF-

15

21 as a Novel Metabolic Regulator," *J. Clin. Invest.* 115: 1627-1635 (2005); Potthoff et al., "FGF21 Induces PGC-
1alpha and Regulates Carbohydrate and Fatty Acid
Metabolism During the Adaptive Starvation Response,"
Proc. Nat'l. Acad. Sci. U.S.A. 106:10853-10858 (2009),
which are hereby incorporated by reference in their entirety)
(FGF21), and phosphate and vitamin D homeostasis (White
et al., "Autosomal Dominant Hypophosphataemic Rickets Is
Associated with Mutations in FGF23," *Nat. Genet.* 26:345-
348 (2000); Shimada et al., "Targeted Ablation of Fgf23
Demonstrates an Essential Physiological Role of FGF23 in
Phosphate and Vitamin D Metabolism," *J. Clin. Invest.*
113:561-568 (2004), which are hereby incorporated by
reference in their entirety) (FGF23). Thus, these ligands
have attracted much attention as potential drugs for the
treatment of various inherited or acquired metabolic dis-
orders (Beenken and Mohammadi, "The FGF Family: Biol-
ogy, Pathophysiology and Therapy," *Nat. Rev. Drug Discov.*
8:235-253 (2009); Beenken and Mohammadi, "The Struc-
tural Biology of the FGF19 Subfamily," in *Endocrine FGFs
and Klothos* (Kuro-o, M. ed.), Landes Bioscience. pp 1-24
(2012), which are hereby incorporated by reference in their
entirety).

Of particular interest is FGF 19, which has been shown to target and have effects on both adipocytes and hepatocytes. For example, mice harboring a FGF19 transgene, despite being on a high-fat diet, show increased metabolic rates, increased lipid oxidation, a lower respiratory quotient and weight loss. Moreover, such mice showed lower serum levels of leptin, insulin, cholesterol and triglycerides, and normal levels of blood glucose despite the high-fat diet and without appetite diminishment (Tomlinson et al., "Transgenic Mice Expressing Human Fibroblast Growth Factor-19 Display Increased Metabolic Rate and Decreased Adiposity," *Endocrinology* 143(5), 1741-1747 (2002), which is hereby incorporated by reference in its entirety). Obese mice that lacked leptin but harbored a FGF19 transgene showed weight loss, lowered cholesterol and triglycerides, and did not develop diabetes. Obese, diabetic mice that lacked leptin, when injected with recombinant human FGF 19, showed reversal of their metabolic characteristics in the form of weight loss and lowered blood glucose (Fu et al., "Fibroblast Growth Factor 19 Increases Metabolic Rate and Reverses Dietary and Leptin-deficient Diabetes," *Endocrinology* 145(6), 2594-2603 (2004), which is hereby incorporated by reference in its entirety).

In one embodiment of the present invention, FGF19 is human FGF19 and has an amino acid sequence of SEQ ID NO: 1 (GenBank Accession No. NP_005108, which is hereby incorporated by reference in its entirety), or a portion thereof, as follows:

sponding to residues spanning residues 1 to 168 of SEQ ID NO:1. In one embodiment, the C-terminal portion of FGF19 begins at a residue corresponding to any one of residues 169, 197, or 204 of SEQ ID NO: 1.

5 In another embodiment, the C-terminal portion of FGF19 of the chimeric protein of the present invention comprises an amino acid sequence spanning residues corresponding to residues selected from the group consisting of from position 204 to 216 of SEQ ID NO: 1, from position 197 to 216 of SEQ ID NO: 1, and from position 169 to 216 of SEQ ID NO:
10 1. In yet another embodiment, the C-terminal portion of FGF19 of the chimeric protein of the present invention comprises an amino acid sequence spanning residues of SEQ ID NO:1, which correspond to residues 191 to 206 or 15 191 to 209 of SEQ ID NO: 100.

In one embodiment of the present invention, FGF19 or a portion thereof is from a mammalian FGF19. In one embodiment of the present invention, FGF19 or a portion thereof is or is from a vertebrate FGF19. In one embodiment, FGF19 or a portion thereof is or is from a non-human vertebrate FGF19. It will be understood that this includes orthologs of human FGF19, or a polypeptide or protein obtained from one species that is the functional counterpart of a polypeptide or protein from a different species. In one embodiment, the C-terminal portion of FGF19 of the chimeric protein of the present invention is from human FGF19. In one embodiment of the present invention, the C-terminal portion of FGF19 is from an ortholog of human FGF19 from *gorilla gorilla*, *pan troglodytes*, *macaca mulatta*, *pongo abelii*, *nomascus leucogenys*, *callithrix jacchus*, *microcebus muri-*
nus, *cholepus hoffmanni*, *ailuropoda melanoleuca*, *sus scrofa*, *bos taurus*, *canis lupus familiaris*, *oryctolagus cuniculus*, *pteropus vampyrus*, *tursiops truncates*, *myotis lucifugus*,
35 *ornithorhynchus anatinus*, *monodelphis domestica*, *anolis carolinensis*, *ochotona princeps*, *cavia porcellus*, *tupaia belangeri*, *rattus norvegicus*, *mus musculus*, *gallus gallus*,
30 *taeniopygia guttata*, *danio rerio*, *xenopus (silurana) tropicalis*, *otolemur garnettii*, *felis catus*, *pelodiscus sinensis*,
35 *latimeria chalumnae*, *mustela putorius faro*, *takifugu rubripes*, *equus caballus*, *oryzias latipes*, *xiphophorus maculatus*, *ictidomys tridecemlineatus*, *gasterosteus aculeatus*, *oreochromis niloticus*, *meleagris gallopavo*, *papio anubis*, *saimiri boliviensis boliviensis*, *pteropus alecto*, *myo-*
40 *tis davidii*, *tupaia chinensis*, or *heterocephalus glaber*.

In other embodiments of the present invention, the portion of FGF19 of the chimeric protein of the present invention is from an ortholog of human FGF19 having an amino acid sequence as shown in Table 1. The portions of an ortholog of human FGF 19 of a chimeric protein according to the present invention include portions corresponding to the

(SEQ ID NO: 1)

1 MRSGCVVVHV WILAGLWLAV AGRPLAFSDA GPHVHYGWGD PIRLRHLYTS GPHGLSSCFL

6 61 RIRADGVVDC ARGQSAHSLL EIKAVALRTV AIKGVHSVRY LCMGADGKMQ GLLQYSEEDC

121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MVPEEPEDLR

181 GHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

In one embodiment, the C-terminal portion of FGF 19 of the chimeric protein of the present invention does not include any of residues 1 to 168 of SEQ ID NO: 1. In certain embodiments of the present invention, the chimeric protein of the present invention does not include residues corre-

65 above-identified amino acid sequences of human FGF 19. Corresponding portions may be determined by, for example, sequence analysis and structural analysis. The high degree of FGF19 sequence conservation among orthologs is shown in FIG. 10.

TABLE 1

Gorilla gorilla (gorilla) FGF19 (Ensembl Accession No. ENSGGOP00000021055, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 2)

1 MRSGCVVVHV WILAGLWLAV AGRPLAFSDA GPHVHYGWD PIRLRHLYTS GPHGLSSCFL
 61 RIRADGVVDC ARGQSAHSLL EIKVALRTV AIKGVHSVRY LCMGADGKMQ GLQYSEEDC
 121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MVPEEPEDLR
 181 GHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Pan troglodytes (chimpanzee) FGF19 (Ensembl Accession No. ENSPTRP00000006877, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 3)

1 MRNGCVVVHV WILAGLWLAV AGRPLAFSDA GRHVHYCWGD PIPLRHLYTS GPHGLSSCFL
 61 RIPANCVVMNC ARGQSAHSLL EIKVALRTV AIKGVHSVRY LCMGADGKMQ GLQYSEEDC
 121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MVPEEPEDLR
 181 GHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Macaca mulatta (Rhesus monkey) FGF19 (GenBank Accession No. XP_001100825, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 4)

1 MRSGCVVVVA WILASLWLAV AGRPLAFSDA GPHVHYGWD PIRLRHLYTS GPHGLSSCFL
 61 RIRTDGVVDC ARGQSAHSLL EIKVALRTV AIKGVHSVRY LCMGADGKMQ GLQYSEEDC
 121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MAPEEPEDLR
 181 GHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Pongo abelii (Sumatran orangutan) FGF19 (GenBank Accession No. XP_002821459, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 5)

1 MRSGCVVVVA WILAGLWLAV AGRPLAFSDS GPHVHYGWD PIRLRHLYTS GPHGLSSCFL
 61 RIRADGVVDC ARGQSAHSLL EIKVALRTV AIKGVHSVRY LCMGADGKMQ GLQYSEEDC
 121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MVPEEPEDLR
 181 RHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Nomascus leucogenys (Northern white-cheeked gibbon) FGF19 (Genbank Accession No. XP_003278071, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 6)

1 MRSECVVVA WILAGLWLAV AGRPLAFSDA GPHVHYGWD PIRLRHLYTS GPHGLSSCFL
 61 RIRADGVVDC ARGQSAHSLL EIKVALRTV AIKGVHSVRY LCMGADGKMQ GLQYSEEDC
 121 AFEEEIRPDG YNVYRSEKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MVPEEPEDLR
 181 GHLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Callithrix jacchus (white-tufted-ear marmoset) FGF19 (GenBank Accession No. XP_002763730, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 7)

1 MWKATAGQQ QOSEAQMSTC PHVPRPLWIA QSCLFSLQLQ YSEEDCAFE EIRPDGYNVY
 61 WSEKHRLPVSS LSSAKQRQLY KKRQFLPLSH FLPMPLPIAPE EPEDLRGHLE SDVFFSPLET
 121 DSMDPFGLVT GLEAVNSPSF EK

Microcebus murinus (mouse lemur) FGF19 (Ensembl Accession No. ENSMICP00000002788, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 8)

1 MPSGQSGCVA ARALILAGLW LTAAGRPLAF SDAGPHVHYG WGEPIRLRHL YTAGPHGLSS
 61 CFLRIRADGS VDCARGOSAH SLLEIRRAVAL RTVAIKGVHSV VRYLCMGADG RMQGLLRYSE
 121 EDCAFEERI PDGYNVYRSE KHRLPVSLSS ARQRQLYKGR CFLPLSHFLP MLPVTPAETG
 181 DLRDHLESMD FASPLETDSM DPFGIATRLG VKSPSFQK

Choloepus hoffmanni (sloth) FGF19 (Ensembl Accession No. ENSCHOP00000002044, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 9) (partial amino acid sequence corresponding to human FGF19 residues 79 to 216)

1 LLEMKAVALR AVAIKGVHSA LYLCMNADGS LHGLPRYSAE DCAFEEIRP DGYNVYWSRK
 61 HGLPVSLSSA KQRQLYKGRG FLPLSHFLP LMPTPAEPAD PGDDVESDMF SSPLETDSMD
 121 PFGIASRLEL VNSPSFQT

Ailuropoda melanoleuca (giant panda) FGF19 (GenBank Accession No. XP_002927952, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 10) (partial amino acid sequence corresponding to human FGF19 residues 12 to 216)

124 VLAGLCL AVAGRPLAFS DAGPHVHYG GEPIRLRHY TAGPHGLSSC FLRIRADGGV
 181 DCARGOSAHS LVEIRRAVAL TVAIKGVHSV RYLCMGADGR MQGLPQYSAG DCAFEEIRP
 241 DGYNVYRSKK HRLPVSLSSA KQRQLYKDRG FLPLSHFLP LPGSPAEPRD LQDHAESDGF
 301 SAPLETDSMD PFGIATKMGV VKSPSFQK

Sus scrofa (pig) FGF19 (Ensembl Accession No. ENSSSCP0000013682, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 11)

1 MRSAPSRCAV VRALVLAGLW LAAAGRPLAF SDAGPHVHYG WGESVRLRHL YTASPHGVSS
 61 CFLRIHSMDGP VDCARGOSAH SLMEIRRAVAL STVAIKGERS RYLCMGADGK MQGQTQYSDE
 121 DCAFEEIRP DGYNVYWSKK HHLPVSLSSA RQRQLYKGRG FLPLSHFLP LSTLPAEPED
 181 LQDPFKSDLF SLPLETDSMD PFRIAALKGA VKSPSFYK

TABLE 1-continued

Bos taurus (bovine) FGF19 (GenBank Accession No. XP_599739, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 12)

136 MRSAP SRCAVARALV LAGLWLAAAG RPLAFSDAGP HVHYGWGESV
 181 RLRHLYTAGP QGLYSCFLRI HSDGAVDCAQ VQSAHSLMEI RAVALSTVAI KGRERSVLYLC
 241 MDADGKMQGL TQYSAEDCAF EEEIRPDGYN VYWSRKHHLP VSLSRRQRQ LFKSRGFLPL
 301 SHFLPMLSTI PAEPEDLQEP LKPDFFLPLK TDSMDPFGLA TKLGSKSPS FYN

Canis lupus familiaris (dog) FGF19 (GenBank Accession No. XP_540802, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 13) (partial amino acid sequence corresponding to human FGF19 residues 25 to 216)

1 LAFSDAGPHV HSFWGPIRL RHLYTAGPHG LSSCFLRIRA DGGVDCARGQ SAHSLMEMRA
 61 VALRTVAIKV VHSGRYLCMG ADGRMQLPQ YSAGDCTE EIRPDGYNVY WSKKKHLPIS
 121 LSSAKQOLY KGRGFLPLSH FLPLPGSPT EPRDLEDHVE SDGFSASLET DSMDPFGIAT
 181 KIGLVVKSPSF QK

Oryctolagus cuniculus (rabbit) FGF19 (GenBank Accession No. XP_002724495, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 14)

1 MRRAPSGAA ARALVLAGLW LAAARPLAL SDAGPHLHYG WGEVRLRHL YATSAHGVSH
 61 CFLRIRADGA VDCERSQSAH SLLEIRAVAL RTVAFKGVS SRYLCMGADG RMRGQLQYSE
 121 EDCAFQEIS SGYNVYRSTT HHLPVSLSSA KQRHLYKTRG FLPLSHFLPV LPLASEETAA
 181 LGDHPEADLF SPPLETDSDM PFGMATKLGP VKSPSFQK

Pteropus vampyrus (megabat) FGF19 (Ensembl Accession No. ENSPVAP00000009339, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 15)

1 MRSPCAVARA LVLAGLWLAS AAGPLALSDA GPHVHYGWGE AIRLRHLYTA GPHGPSSCFL
 61 RIRADGAVDC ARGQSAHSLV EIRAVALRNV AIKGVHSVRY LCMGADGRML GLQYSADD
 121 EDCAFEER PDGYNVYWSK KHHLPVSLSS ARQRLQFKGR GFPLSHFLP RSPTEPENFE
 181 DHLEADTFSS LETDDMDPFG IASKLGLEES PSFQK

Tursiops truncatus (dolphin) FGF19 (Ensembl Accession No. ENSTTRP00000000061, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 16)

1 MRSAPSRCAV ARALVLAGLW LAAAGRPLAF SDAGPHVHYG WGESVRLRHL YTAGPQGLSS
 61 CFLRIRADGA VDCAPVQSAH SLLEIRAVAL STVAIKGERS VLYLCMGADG KMQGLSQYSA
 121 EDCAFEER PDGYNVYWSK KHHLPVSLSS ARQRLQFKGR GFPLSHFLP MLSTIPTEPD
 181 EIQDHLPKPD LALPKTDSDM DPFLGLATKLG VVKSPSFQK

Myotis lucifugus (microbat) FGF19 (Ensembl Accession No. ENSMLUPO0000002279, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 17)

1 MQSAWSRRV ARALVLASLG LASAGGPLGL SDAGPHVHYG WGESIRLRHL YTSGPHGPSS
 61 CFLRIRADGA VDCAPVQSAH SLLEIRAVAL RKVAIKGVHS ALYLCMGGDG RMLGLPQFSP
 121 EDCAFEER PDGYNVYRSQ KHQLPVSLSS ARQRLQFKAR GFPLSHFLP MLPSSPAGPV
 181 PRERPSEPDE FSSPLETDSDM DPFGIANNLR LVRSPSFQK

Ornithorhynchus anatinus (platypus) FGF19 (GenBank Accession No. XP_001506714, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 18) (partial amino acid sequence corresponding to human FGF19 residues 79 to 216)

1 MLSCVVLPSL LEIKAVAVRT VAIKGVHISR YLCMEEDGKT PWARLLEIKA VAVRTVAIKG
 61 VHSSRYLCME EDGKLHGQIW YSAEDCAFEE EIRPDGYNVY KSKKYGVPS LSAAKQRQQF
 121 KGRDFLPLSR FLPMINTV PVPAEGFDYAD YFESDIFSSP LETDSMDPFR IAPKLSPVKS
 181 PSFQK

Monodelphis domestica (opossum) FGF19 (GenBank Accession No. XP_001506714, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 19)

1 MAQLLAPLLT LAALWLAPTA RARPLVDAGP HVYYGWGEPI RLRHLYTANR HGLASFSLR
 61 IHRDGRVGDGS RSQSAISLLE IKAVALRMVA IKGVHSSRYL CMGDAGKLOG SVRFAEDCT
 121 FEEQIRPDGY NVYQSPKYNL PVSLCTDKQR QQAHGKEHLP LSHFLPMINA IPLEAEPEG
 181 PRMLAAPLET DSMDPFLGLTS KLLPVKSPSF QK

Anolis carolinensis (anole lizard) FGF19 (GenBank Accession No. XP_003214715, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 20)

1 MCRRALPPLL ALLGLAAVAS RALPLTDAGP HVSYGWGEPI RLRHLYTAGR QGLFSQFLRI
 61 HADGRVGDAG SQNRQSLLEI RAVSLRAVAL KGKVHSSRYLC MEEDGRLRG MRYSAEDCSF
 121 EEEMRPDGYN IYKSKKYGVL VSLSNARQRQ QFKGKDFLPL SHFLPMINTV PVESADFGEY
 181 GDTRQHYESD IFSSRLETDS MDPFGLTSEV SSVQSPSFQK

Ochotona princeps (pika) FGF19 (Ensembl Accession No. ENSPOPR00000009838, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 21) (partial amino acid sequence corresponding to human FGF19 residues 12 to 77 and 113 to 216)

1 VRSRGAMARA LVLATLWLA GRPLALSDA GPHLHYGWGE PIRLRHLYAT SAHGLSHCFL
 61 RIRTDGTVC ERSQSAH----- ----- ----- ----- LQYSEEDC

TABLE 1-continued

121 AFEEEIISGY NVYRSRRYQL PVSLGSARQR QLQRSGRFLP LSHFLPVLP ASEEVAAPAD
 181 HPQADPFSPM ETDSMDPFGM ATKRLVKSP SFQK

Cavia porcellus (guinea pig) FGF19 (Ensembl Accession No. ENSCPOP00000007325, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 22)

1 MWSAPSGCVV IRALVLAGLW LAVAGRPLAR RSLALSDQGP HLYYGDQPI RLRHLYAAGP
 61 YGRSRCLRLI HTDGAVDCVE EQSEHCCLLEI RAVALETVAI KDINSVRYLC MGPDGRMRGL
 121 PWYSEEDCAF KEEISYPGYS VYRSOKHHLP IVLSSVKQRQ QYOSKGVVPL SYFLPMLPKA
 181 SVEPSDEEES SVFSLPLKTD SMDFPGMASE IGLVKSPSFQ K

Tupaia belangeri (tree shrew) FGF19 (Ensembl Accession No. ENSTBEP00000000264, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 23) (partial amino acid sequence corresponding to human FGF19 (residues 1 to 112 and 136 to 216))

1 MRRTPSGFAV ARVLEPLGSLW LAAAGSPLAL SDAGPHVNNG WDESIRLRHL YTASPHGSTS
 61 CFLRIRDGS VDCARGQSLH SLLEIKAVAL QTVAIKGVYS VRYLCMDADG RMQGL-----
 121 ----- ST KHGLPVSLSS AKQRQLLTVR GFPSLPHFLL MMAKTSAGPG
 181 NPRDHPGSNT FSLPLETDSM DPFGMTTRHG LVKSPSFQN

Rattus norvegicus (Norway rat) FGF15 (GenBank Accession No. NP_570109, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 24)

1 MARKWGRIV ARALVLATLW LAVSGRPLVQ QSQSVSDEGP LFLYGWGKIT RLQYLYSAGP
 61 YVSNCFLRIR SDGSVDCEED QNERNLLEFR AVALKTIAIK DVSSVRYLCM SADGKIYGLI
 121 RYSEEDCTFR EEMDCLGYNQ YRSMKHHHLHI IFIKAQPREQ LQGQKPSNFI PIFHRSFFES
 181 TDQLRSKMF5 LPLESDSMDP FRMVVEDVDHL VKSPSFQK

Mus musculus (house mouse) FGF15 (GenBank Accession No. NP_032029, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 25)

1 MARKWNGRAV ARALVLATLW LAVSGRPLAQ QSQSVSDEDP LFLYGWGKIT RLQYLYSAGP
 61 YVSNCFLRIR SDGSVDCEED QNERNLLEFR AVALKTIAIK DVSSVRYLCM SADGKIYGLI
 121 RYSEEDCTFR EEMDCLGYNQ YRSMKHHHLHI IFIKAQPREQ LQDQKPSNFI PIFHRSFFES
 181 GDQLRSKMF5 LPLESDSMDP FRMVVEDVDHL VKSPSFQK

Gallus gallus (chicken) FGF19 (GenBank Accession No. NP_990005, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 26)

1 MGPAPPAAPG AALALLGIAA AAAARASLPL PDVGGPHVNY GWGEPIRLRH LLHRPGKHGL
 61 FSCFLRIGGD GRVDAVGQS PQSLEIRAV AVRVAIKGV QSSRYLCMDE AGRLHQLSY
 121 SIEDCSFEEE IRPDGYNVYK SKKYGIVSL SSAKQRQFK GKDFLPLSHF LPMINTVPVE
 181 VTDFGEYGDY SQAFEPEVYS SPLETDSDMP FGITSKLSPV KSPSFQK

Taeniopygia guttata (zebra finch) FGF19 (GenBank Accession No. XP_002194493, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 27)

1 MVIISNLYLQ QNDVMNMRR APLRVHAARS SATPASALPL PPPDAGPHLK YWGEPRLR
 61 HLYTASKHGL FSCFLRIGAD GRVDAAGSQS PQSLEIRAV AVRVAIKGV QSSRYLCMDE
 121 AGRLHQCLRN STEDCSFEEE IRPDGYNVYK SKKYGIVSL SSAKQRQFK GKDFLPLSHF
 181 LPMINTVPME SADFGEYGDY SQAFEAEAFS SPLETDSDMP FGIAASKLSLV KSPSFQK

Danio rerio (zebrafish) FGF19 (GenBank Accession No. NP_001012246, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 28)

1 MLLLLFVTVC GSIGVESLPL PDSPGPHLAND WSEAVRLRHL YAARHGLHQ INTDGEIIGS
 61 TCKARTVSLM EIWPVDTGCV AIKGVASSRF LCMERLGNLY GSHIYTKEDC SFLERILPDG
 121 YNVYFSSKG ALVTLGAKN KLHSNDGTSQ SQFLPMINTL SEEHTKQHSG EQHSSVNHGQ
 181 DHQGLEIDS MDPFGKISQI VIQSPSFNKR

Xenopus (Silurana) tropicalis (western clawed frog) FGF19 (GenBank Accession No. NP_001136297, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 29)

1 MWKTLFWILV PMMVAVLYFL GGAESPLFLD AGPHMQNGWG ESIRIRHLYT ARRFGHDSSY
 61 LRIHEDGRVD GDRQOSMHSL LEIRAIAVGI VAIKGYRSSL YLCMGSEGKL YMGMHSYSQDD
 121 CSFEEELLPD GYNMYKSRKH GVAVSLSKBK QKQQYKGKGY LPLSHFLPVI SWVPMEPTGD
 181 VEDDIYRFPF NTDTKSVIDS LDTLGLMDFS SYHKK

Otolemur garnettii (bushbaby) FGF19 (Ensembl Accession No. ENSOGAP00000017975, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 30)

1 MPSGLRGRVV AGALALASFW LAVAGRPLAF SDAGPHVHYG WGEPIRLRHL YTAGPHGLSS
 61 CFLRVRTDGA VDCARGQSAH SLLEIRAVAL RTVAIKGVHS ARYLCMGADG RMQGLPQYSE
 121 EDCAFEERI PDGYNVYWSK KHRLPVSLSS ARQQLYKGR GFLPLSHFLP MLPVTPAEPG
 181 DLRDHLKESDM FSLPLETDSM DPFGIATRLG VVKSPSFQK

Felis catus (cat) FGF19 (Ensembl Accession No. ENSFCAP0000022548, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 31)

1 MRSAPSQCAY TRALVLAGLW LAAAGRPLAF SDAGPHVHYG WGEPIRLRHL YTAGPHGLSS
 61 CFLRIRADGG VDCARSQSAH SLVEIRAVAL RTVAIKGVHS VRYLCMGADG RMQGLLQYSA

TABLE 1-continued

121 GDCAFQEEIR PDGYNVYRSE KHRLPVSLS AIQRQLYKGR GFLPLSHFLP MLPGSPAEP
 181 DLQDHVESER FSSPLETDSM DPFGIATKMG LVKSPSFQK

Pelodiscus sinensis (Chinese softshell turtle) FGF19 (Ensembl Accession No. ENSPSIP00000010374, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 32)

1 MWRSLCKSHT SLALLGLCFA VVRSLPLFSD AGPHVNNGWG EPIRLRHYT ASRHGLFNYF
 61 LRISSDGKVD GTSIQSPHSL LEIRAVAVRT VAIKGVHSSR YLCMEEDGKL HGLLRYSTED
 121 CSFEEIRPD GYNVYKSKKY GISVSLSSAK QRQFKGKDF LPLSHFLPMI NTVPVESMDF
 181 GEYGDYSHTF ESDLFSSPLE TDSDMPFGIT SKISPVKSPS FQK

Latimeria chalumnae (coelacanth) FGF19 (Ensembl Accession No. ENSLACP00000014596, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 33)

1 MLQALYNLCT ALVLFKLPFA MVGYTLPAN EGPHLNWDWG ESVRLKHYT SSKHGLISYF
 61 LQINDDGKVD GTTTRSCYSL LETIKSVGPVG LAIKGIQSSR YLCVEKDGL HGSRTYSADD
 121 CSFKEDILPD GYTIYVSKKH GSIVNLSNHK QRQRNRRTL PPFSQFLPLM DTIRVECMNC
 181 GEHCDDNLHD ELETGLSMDP FESTSKKSPQ SPSFHNR

Mustela putorius furo (ferret) FGF19 (Ensembl Accession No. ENSMPUP00000004571, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 34)

1 MRSAAASRCAV ARALVLAGLW LAAAGRPLAF SDAGPHVHYG WGEPIRLRHL YTAGPHGLSS
 61 CFLRIRADGG VDCARGOSAH SLVEIRAVAL RTVAIKGVYS DRYLCMGADG RMQGLPQYSA
 121 GDCAFEEEIR PDGYNVYRSLK KHRLPVSLS AKQRQLYKDR GFLPLSHFLP MLPGSLAEPR
 181 DLQDHVEADG FSAPLETDSM DPFGIATKMG LVKSPSFQK

Takifugu rubripes (fugu) FGF19 (Ensembl Accession No. ENSTRUP00000007110, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 35)

1 SSTRISGNMV LLMLLPITVAN LFLCAVGVLSL PLLDQGSHFP QGWEQVVRFR HLYAASAGLH
 61 LLITEEGSIQ GSADPTLYSL MEIRPVDPGC VVIRGAATTR FLCIEGAGRL YSSQTYSKDD
 121 CTFREQILAD GYSVYRSVGH GALVSLGNR QQLRGEDWSV PTLAQFLPRI SSLDQDFKAA
 181 LDETEKPEQT APQRSEPVDM VDSFGKLSQI IHSPSFHK

Equis caballus (horse) FGF19 (Ensembl Accession No. ENSECAP00000017705, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 36); partial sequence corresponding to human FGF19 residues 20 to 113

1 AAGRPLALSD AGPHVHYGWG EPIRLRHYT AGPHGLSSCF LRIRADGAVD CARGQSAHSL
 61 VEIRAVALRT VAIKGVHSVR YLCMGADGRM QCLV

Oryzias latipes (medaka) FGF19 (Ensembl Accession No. ENSORLP00000000352, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 37)

1 TMLLIVVTIS TMVFSDSGVCS SMPLSDHGPH ITHWSQVVR LRHLYAVKPG QHVQIREDGH
 61 IHGSABQTLN SLLEIRPVAP GRVVFVRGVAT SRFLCMESDG RLFFSHTFDK DNCVPREQIL
 121 ADGYNIIYISD QHGTLLSLGN HRQRQGLDR DVPAQFLP RISTLQQGVY PVPDPPHQMR
 181 TMQTEKTLDA TDTFGQLSKI IHSPSFNK

Xiphophorus maculatus (platyfish) FGF19 (Ensembl Accession No. ENSXMAP00000001516, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 38)

1 MFVFILCIAG ELFTLGVFCM PMMDQGPLVT HGWQVVRHR HLYAAKPGHL LLISEDGQIH
 61 GSADQTLYSL LEIQPVGPGR VVIKGVATR FLCMESDGRL YSTETYSRAD CTFREQIQAD
 121 GYNVYTDSH GALLSLGNQ QRHSGSDRGV PALARFLPRL NTLQQAVPTE PDVPDQLSPE
 181 KVQQTVDMVA SFGKLSHIIH SPSFHNR

Ictidomys tridecemlineatus (squirrel) FGF19 (Ensembl Accession No. ENSSTOP00000021639, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 39)

1 MRSAPSPLAR ARALVLASLW LAVAGRPLAR RSLALSDQGP HLYYGWDQPI RLRLHYAAGP
 61 YGFSNCFLRI RTDGAVDCEE KQSERSLMEI RAVALETVAI KDINSVRYLC MGADGRIQGL
 121 PRYSEEECTF KEEISYDGYN VYRSQKYHLP VVLSAKQRQ LYQSKGVVPL SYFLPMLPLA
 181 SAETRDRLES DVFSLPLETD SMDFPGMASE VGLKSPSFQK

Gasterosteus aculeatus (stickleback) FGF19 (Ensembl Accession No. ENSGACP00000018732, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 40)

1 MLLLLPVAYV ASVFLALGVV CLPLTDQGLH MADDWGQSVR LKHLYAASPG LHLLIGEDGR
 61 IQGSAQSPY SLLEISAVDP GCVVIRGVAT ARFLCIEGDG RLYSSDTYSR DDCTFREQIL
 121 PDGYSVYVSH GHGALLSLGN HRQRQGRDH GVPALAQFLP RVSTMQASA PDAPGQTAPE
 181 TEEPVDSFGK LSQIIHSPSF HER

Oreochromis niloticus (tilapia) FGF19 (Ensembl Accession No. ENSONIP00000022796, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 41)

1 MLLLLIVSIV NMLFGVGMVC MPLSDNGPHI AHGWAQVVR RLHYATRPGM HLLISEGGQI
 61 RGSAVQTLHS LMEIRPVGPB RVVIRGVATA RFLCIEDDGT LYSSHAYSRE DCIFREQILP

TABLE 1-continued

121 DGYNIYISDR HGVLSSLGHN RQRLQGLDRG DPALAQFLPR ISTLNQIPSP GANIGDHMKV
 181 AKTEEPVDTI DSFGKFSQII DSPSFHKR

Meleagris gallopavo (turkey) FGF19 (Ensembl Accession No. ENSMGAP00000010265, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 42); partial sequence corresponding to human FGF19 residues 71 to 216
 1 VGNQSPS01 EITAVDVGIV AIKGLFSGRY LAMNKGRRLY ASLSYSIEDC SFEEEIRPDG
 61 YNVYKSKKYG ISVSLSSAKQ RQPFKGKDFL PLSHFLPMIN TVPVEVTDFG EYGDYSQAFE
 121 PEVYSSPLET DSMDPFGITS KLSPVKSPSF QK

Papio anubis (olive baboon) FGF19 (GenBank Accession No. XP_003909471, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 43)
 1 MRSGCCVVVA WILASLWLAV AGRPLAFSDA GPHVHYGWGD PIRLRHLYTS GPHGLSSCFL
 61 RIRTDGVVDC ARGQSAHSLL EIKAVALRTV AIKGVHSVRY LCMGADGKMQ GLLQYSEEDC
 121 AFEEEIRPDG YNVYRSQKHR LPVSLSSAKQ RQLYKNRGFL PLSHFLPMLP MAPEEPEDLR
 181 GPLESDMFSS PLETDSMDPF GLVTGLEAVR SPSFEK

Saimiri boliviensis boliviensis (Bolivian squirrel monkey) FGF19 (GenBank Accession No. XP_003941214, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 44)
 1 MRSGCCVVVA WILAGLWLAV VGRPLAFSDA GPHVHYGWGD PIRLRHLYTS SPHGLSSCFL
 61 RIRSDGVVDC ARGQSAHSLL EIKAVALRTV AIKGVHSRYY LCMGADGRLQ GLFQYSEEDC
 121 AFEEEIRPDG YNVYLSEKHH LPVSLSSAKQ RQLYKDRGFL PLSHFLPMLP RAPEEPDDLR
 181 GHLESDVFSS PLETDSMDPF GLVTGLEAVN SPSFEK

Pteropus alecto (black flying fox) FGF19 (GenBank Accession No. ELK13233, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 45)
 1 MRSPCAVARA LVLAGLWLAS AAGPLALSDA GPHVHYGWGE AIRLRHLYTA GPHGPSSCFL
 61 RIRADGAVDC ARGQSAHSLV EIRAVALRNV AIKGVHSVRY LCMGADGRML GLLQYSADD
 121 AFEEEIRPDG YNVYHSKKHH LPVSLSSAKQ RQLYKDRGFL PLSHFLPMLP RSPTEPENFE
 181 DHLEADTFSS PLETDSMDPF GIASKLGLEE SPSFQK

Myotis davidii (David's myotis) FGF19 (GenBank Accession No. ELK24234, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 46)
 1 MSGQNSGRHG SRPGLDEEPE PGPLERLALG STRADPQLCD FLENHFLGYT CLELDICLAT
 61 YLGVSHWGES IRLRHLYTSG PHGPSCFLR IRVDGAVDCA RGQSAHSLVE IRAVALRKVA
 121 IKGVHSALYL CMEGDGRMRG LPQFSPEDCA FEEEIRPDGY NVYRSQKHQL PVSLSSARQR
 181 QLFKARGFLP LSHFLPMLPS SPAEPVHRER PLEPDAFSSP LETDSMDPFG IANNLRLVKS
 241 PSFQK

Tupaia chinensis (Chinese tree shrew) FGF19 (GenBank Accession No. ELW64990, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 47); residues 1-257, excluding 13-19
 1 MRRTWSGFAV AT-----R AGSPLALADA GPHVNYGWDE SIRLRHLYTA SLHGSTSCFL
 61 RIRDDGSVGC ARGQSMHSLL EIKAVALQTV AIKGVYSVRY LCMTDGRM QLPOYSEEDC
 121 TFEEEIRSDG HNVYRSKKHG LPVSLSSAKQ RQLYKGRGFL SLSHFLMMMP KTSAGPGNPR
 181 DQRNPDRQD PNTFLSLPLET DSMDPFGMTT RHGLLLDSCC ASLVLLNIST DGEFSPYGN
 241 LRPSFRFKLF KMKKVTN

Heterocephalus glaber (naked mole-rat) FGF19 (GenBank Accession No. EHB12332, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 48)
 1 MRFSKSTCGF FNHQRLQALW LSLSSVKWVL DAAVEGRPIR LRHLYAAGPY GRSRCFLRIH
 61 TDGAVDCVVE QSEHCLLEIR AVALETAIAK DINSVRYLCM GDPGRMQGLP WYSEEDCAFK
 121 EEISYPGYSV YRSQKHHLPY VLSSVKQRQQ YQSKGVVPLS YFLPMLPKAS VEPGDEEESA
 181 FSLPLKTDSM DPFGMASEIG LAKSPSFQK

In one embodiment, a C-terminal portion of FGF19 of the chimeric protein of the present invention comprises the conserved amino acid sequence TGLEAV(R/N)SPSFEK (SEQ ID NO: 49). In one embodiment, a C-terminal portion of FGF19 comprises the conserved amino acid sequence MDPFGLVTGLEAV(R/N)SPSFEK (SEQ ID NO: 50). In one embodiment, the C-terminal portion of FGF19 of the chimeric protein of the present invention comprises the conserved amino acid sequence LP(M/I)(V/A)PEEPEDLR (G/R)HLESD(M/V)FSSPLETDSDMPFGLVTLGAEV(R/N)SPSFEK (SEQ ID NO: 51).

In one embodiment, the C-terminal portion of FGF 19 of the chimeric protein of the present invention consists of an amino acid sequence selected from the group consisting of TGLEAV(R/N)SPSFEK (SEQ ID NO: 49); MDPF-

GLVTGLEAV(R/N)SPSFEK (SEQ ID NO: 50); and LP(M/I)(V/A)PEEPEDLR(G/R)HLESD(M/V)FSSPLETDSDMPFGLVTLGAEV(R/N)SPSFEK (SEQ ID NO: 51).

In certain embodiments according to the present invention, the C-terminal portion of FGF19 of the chimeric protein of the present invention includes a polypeptide sequence that has at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% amino acid sequence identity to the amino acid sequences of any of SEQ ID NOS: 49 to 51. In certain embodiments according to the present invention, the C-terminal portion of FGF19 of the chimeric protein of the present invention includes a polypeptide sequence that has at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% amino acid

sequence homology to the amino acid sequences of any of SEQ ID NOs: 49 to 51.

Percent (%) amino acid sequence identity with respect to a given polypeptide sequence identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical to the amino acid residues in the reference sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent (%) amino acid sequence homology with respect to a given polypeptide sequence identified herein is the percentage of amino acid residues in a candidate sequence that are identical to or strongly similar to the amino acid residues in the reference sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence homology. Strongly similar amino acid residues may include, for example, conservative amino acid substi-

tutions known in the art. Alignment for purposes of determining percent amino acid sequence identity and/or homology can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared.

It will be understood that the portion from FGF19 of the chimeric protein of the present invention may be from a nucleotide sequence that encodes an FGF19 protein (e.g., those encoding orthologs) from a mammal or even a non-mammalian species. For example, a nucleotide sequence encoding a mammalian or non-mammalian FGF19 protein according to the present invention may include, but is not limited to, those FGF-encoding nucleotide sequences shown in Table 2.

TABLE 2

Human FGF19 gene coding sequence (SEQ ID NO: 52) (GenBank Accession No. NM_005117, which is hereby incorporated by reference in its entirety)

```

464 ATGCGGA CGGGGTGTGT GTGGTGCAC GTATGGATCC TGGCCGGCTC CTGGCTGGCC
521 GTGGCCGGC CCCCCCTCGC CTTCTCGGAC GCGGGGCCCG ACAGTGACTA CGGCTGGGGC
581 GACCCCATCC GCGTCGGCA CCTGTACACC TCCGGCCCCC ACGGGCTCTC CAGCTGCTC
641 CTGCGCATCC CGTGCAGACGG CGTCGCTGGC TGGCGCGGG GCGAGACGGC GCACAGTTG
701 CTGGAGATCA CGGAGCTCGC TGCGGACCC GTGCGCATCA AGGGCGTCA CAGCGTCCG
761 TACCTCTGCA TGGGGCCCGA CGGAAGATC CAGGGCTTC TTCACTACTC GGAGGAAGAC
821 TGTGCTTTCG AGGAGGAGAT CGGCCAGAT GGCTACAATG TGTACCGATC CGAGAACGAC
881 GCCTCCCGG TCTCCCTGAG CAGTGCCTAA CAGGGCGACG TGTACAAGAA CAGAGGCTT
941 CTTCACACTC CTCACTCTC CCCATGCTCC CCCATGGTCC CAGAGGAGCC TGAGGACCTC
1001 AGGGGCCACT TGGAAATCTGA CATGTTCTCT CGCCCGCTGG AGACCGACAG CATGGACCCA
1061 TTGGGCTTG TCACCGGACT GGAGGCCGTG AGGAGTCCCA GCTTGAGAAA GTAA

```

Gorilla FGF19 gene coding sequence (SEQ ID NO: 53) (Ensembl Accession No. ENSGGOT00000028361, which is hereby incorporated by reference in its entirety)

```

463 ATGCGGAG CGGGGTGTGT GTGGTGCACCG TCTGGATCTC GGCGGGCTC TGGCTGGCCG
521 TGGCCGGCG CCCCCCTCGC TTCTCGGAC CGGGGCCCGA CGTGCACACTAC GGCTGGGGCG
581 ACCCCATCCG CCTCGGGCAC CTGTACACCCT CGGGCCCCCA CGGCTCTCC AGCTGTTCC
641 TGCGCATCCG TGCGCAGCGC GTCTGGACT CGCGCGGGGG CCAGAGCGCG CACAGTTGC
701 TGGAGATCAA CGCAGTGTGT CTGGGACCC CGGGCATCAA CGGGCGTCA CAGCGTCCGGT
761 ACCTCTGCA CGGCGCCGAC CGCAAGATGC AGGGGCTGTG TCAGTACTCG GAGGAAGACT
821 GTGCTTTGCA CGAGGAGATC CGGCCAGATG GTACAATGTT GTACCGATCT GAGAACGACC
881 GCCTCCCGGT CTCCCTGAGC AGTGCCTAAAC AGCGGCAGCT GTACAAGAAC AGAGGCTT
941 TTCCGCTCTC CAATTTCTGCCC CCCATGCTGC CCATGGTCC AGAGGAGCCT GAGGACCTCA
1001 GGGGCCACTT GGAATCTGAC ATGTTCTCT CACCCCTGGA GACCGACAGC ATGGACCCAT
1061 TTGGGCTTGT CACCGGACTG GAGGCCGTG GGAGTCTAG CTTTGAGAAG TAA

```

Pan troglodytes gene coding sequence (chimpanzee) FGF19 (SEQ ID NO: 54) (Ensembl Accession No. ENSPTRT00000007454, which is hereby incorporated by reference in its entirety)

```

1 ATGCGGAACG GGTGTGTGGT GGTCCACGTC TGGATCCTGG CGGGCCTCTG GCTGGCCGTG
61 GCGGGGCCCG CCCTCGGCC TTCTCGGACCG GGGGCCAACG TGCACTACTG CTGGGGGCCGAC
121 CCCATCCCCC TGCGGACCTC GTACACCTCC CGGGCCCCATG GGCTCTCCAG CTGCTCTCTG
181 CGGATCTCTG CGAACCTGCGT CATGAACTGCG CGCGGGGGCC AGAGCGCGCA CAGTTTGCTG
241 GAGATCAAGG CAGTCGCTCT CGGGCACCGT GGCATCAAGG CGCTGCACAG CGTGCCTGAC
301 CTCTGCGATGG CGGCGCACGG CAAGATGCGAG GGGCTGCTTC AGTACTCGA GGAAGACTGT
361 GCTTTCGAGG AGGAGATCCG CCCAGATGCC TACAATGTT ACCGATCCGA GAGGCACCGC
421 CTCCCGGTCT CGCTGAGCAG CGGCCAGCTGT ACAAGAACAG AGGCTTCTT
481 CCACTCTCTC ATTCTCTGCC CATGCTGCC ATGTTCCCG AGGAGCTGAGA GGACCTCAGG
541 GCCCACTGG AATCTGACAT GTTCTCTCTG CCCCTGGAGA CCGACAGCAT GGACCCATT
601 GGCGTTGCA CGGGACTGGG AGTCCCAGT TTGAGAAGTA A

```

Macaca mulatta gene coding sequence (Rhesus monkey) FGF19 (SEQ ID NO: 55) (GenBank Accession No. XM_001100825, which is hereby incorporated by reference in its entirety)

```

758 ATG AGGAGCGGGT GTGTGGTGGT CCACGCCCTGG ATCCTGGCCA GCCTCTGGCT
811 GGCCTGTGCC CGGCGTCCCC TCGCTTCTC GGACGCCGGG CCCCACGTGC ACTACGGCTG
871 GGGGACCCCC ATCCGCTCGC GGACACCTGTA CACCTCCGGC CCCCATGGGC TCTCCAGCTG
931 CTTCCTGGC ATCCGACCG ACGGCGTCTG GGACTGCGG CGGGGCAAA CGCGCACAG
991 TTGCTGGAG ATCAAGGCAG TAGCTCTGCG GACCGTGGCC ATCAAGGGCG TGACAGCGT
1051 GCGGTACCTC TGCATGGCG CCGACGGCAA GATGCGAGGG CTGCTTCAGT ACTCAGAGGA
1111 AGACTGTGCT TTGAGGAGG AGATCCGCC TGATGGCTAC AATGTATAAC GATCCGAGAA
1171 GCACCGCCTC CGGGTCTCTC TGAGCAGTGC CAAACAGAGG CAGCTGTACA AGAACAGAGG
1231 CTTCTCTCCG CTCTCTCATT TCCTACCCAT GCTGCCCATG GCGCCAGAGG AGCCTGAGGA

```

TABLE 2-continued

1291 CCTCAGGGGC CACTTGAAT CTGACATGTT CTCTTCGCC CTGGAGACTG ACAGCATGGA
 1351 CCCATTGGG CTTGTACCG GACTGGAGC GGTGAGGAGT CCCAGCTTG AGAAATAA

Pongo abelli gene coding sequence (Sumatran orangutan) FGF19 (SEQ ID NO: 56) (GenBank Accession No. XM_002821413, which is hereby incorporated by reference in its entirety)

763 ATGCGGAG CGGGTGTGTG GTGGTCCACG CCTGGATCCT GGCGGGCCTC TGGCTGGCG
 821 TGCCCGGGCG CCCCTGGCC TTCTCGACT CGGGGCCCCA CGTGCACCTAC GGCTGGGGCG
 881 ACCCCATCC CTCGCGCAC CTGTCACCT CGGGCCCCCA CGGCTCTCC AGCTGCTTCC
 941 TGCGCATCG TGCGCGACG GTCGTGACT CGCGCGGGGG CCAGAGCGCG CACAGTTGC
 1001 TGGAGATCAA GGCAGTCGCT CTGCGGACCG TGCCATCAA GGGCGTGCAC AGCGTGCAG
 1061 ACCTCTGCA GGGCGCCGAC GGCAAGATGC AGGGGCTGCT TCAGTACTCG GAGGAAGACT
 1121 GTGCTTCGA GGAGGAGATC CGCCCGAGATG GCTACAATGT GTACCGATCC GAGAACGACC
 1181 GCCTCCGGT CTCCCTGAGC ATGCGAAAC AGCGCGAGCT GTACAAGAAC AGGGGCTTC
 1241 TTCCGCTCTC TCATTCCTG CCCATGCTGC CCATGGTCCC AGAGGAGCCT GAGGACCTA
 1301 GGCGCACTT GGAATCCGAC ATGTTCTCTT CGGCCCTGGA GACCGACAGC ATGGACCCAT
 1361 TTGGGCTTGT CACCGGACTG GAGGCCGTGA GGAGTCCCAG CTTTGAGAAA TAA

Nomascus leucogenys gene coding sequence (Northern white-cheeked gibbon) FGF19 (SEQ ID NO: 57) (GenBank Accession No. XM_003278023, which is hereby incorporated by reference in its entirety)

456 ATGGC GAGCGACTGT GTGGTGTCC ACAGCTGGAT CCTGGCCGGC CTCTGGCTGG
 511 CAGTGGCCGG GCGCCCCCTC GCCTTTCCG ACAGGGGGCC CCACGTGCAC TACGGCTGG
 571 GCGACCCAT CGCTCTGCGG CACCTGTACA CCTCCGGCC CGACGGGCTC TCCAGCTGCT
 631 TCTCGCGCAT CGCTGCGGAC GGCCTGCTGG ACTCGCGCGC GGGCGCAGAGC CGCGCACAGT
 691 TGCTGGAGAT CAAGGGCAGTC GCTCTGCGGA CCCTGCGCAT AAAGGGCGTG CACAGCGTGC
 751 GGTACCTCTG CATGGCGCC GACCGCAAGA TGCAAGGGCT GCTTCAGTAT TCGGAGGAAG
 811 ACTGTGCTTT CGAGGAGGAG ATCCGCCCAG ATGGCTACAA TGTTGACCGA TCCGAGAACG
 871 ACCGCTCCCTG CGCTCTCCCG AGCACTGCGCA AACAGCGCGA GCTGTATAAG AACAGAGGCT
 931 TTCTTCCACT CTCTCATTC CTGGCCATGC TGCCCATGGT CCCAGAGGAG CCTGAGGACC
 991 TCAGGGGCA CTTGGAATCT GACATGTTCT CTTCGCCCCG GGAGACCGAC AGCATGGACC
 1051 CATTGGGCT TGTCACCGGA CTGGAGGCCG TGAGGAGTCC CAGCTTTGAG AAATAA

Callithrix jacchus gene coding sequence (white-tufted-eared marmoset) FGF19 (SEQ ID NO: 58) (GenBank Accession No. XM_002763684, which is hereby incorporated by reference in its entirety)

1 ATGTGGAAGG CCACCGCTGG TGGCCAGCAG GGACAGTCGG AAGCACAAAT GTCCACATGT
 61 CCCCATGTT CTCGCTCT GTGGATTGCT CAGAGCTGCT TGTTTCTCT GCAGCTCCAG
 121 TACTCGAGG AAGACTGTGC TTTCAGGAG GAGATCCGCC CTGATGGCTA CAATGTGTAC
 181 TGGCCGAGA AGCACCGCT CCCGGCTCCTC CTGAGCAGCG CCAAACAGCG GCAGCTGTAC
 241 AAAGAACGAG GCTTCTCTTCTG ACTGTCCCAT TTCTGCCCCA TGCTGCCAT AGCCCCAGAA
 301 GAGCCTGAG ACCTCAGGG ACACCTGGAA TCTGACGTGT TCTCTTCACC CCTGGAGACT
 361 GACAGCATGG ACCCATTTGG GCTTGTACCG GGACTGGAGG CGGTGAACAG TCCCAGCTT
 421 GAGAAGTAA

Microcebus murinus gene coding sequence (mouse lemur) FGF19 (SEQ ID NO: 59) (Ensembl Accession No. ENSMICT00000003065, which is hereby incorporated by reference in its entirety)

1 ATGCCGAGCG GCGAAAGCGG TTGTGCGGC GCGCGCGCC TGATCCTGGC CGGCCTCTGG
 61 CTGACCGGG CGGGCGCCC GCTGGCTTC TCCGACCGG GCGCGCACGT GCACACTCGC
 121 TGGGGCGAGC CCATCCGCTC CGGGCACCTG TACACCGCGG GCGCCCGACGG CCTCTCCAGC
 181 TGCTTCTGC GCATCCGCG AGACGGCTCC GTGGACTGCG CGCGGGGCCA GAGCGCACAC
 241 AGTTGCTGG AGATCAGGGC GGTGCGCTCTT CGGACTGTGG CCATCAAGGG CGTGCACAGC
 301 GTGCGGTAC TCTGCGATGG CGCAGACGGC AGGATGCGAGG GGCTGCTCCG GTACTCGGAG
 361 GAAGACTGTG CCTCGAGGA GGAGATCCGC CCCGATGGCT ACACAGTGTAA CGGGCTTGAG
 421 AAGCACCGCC TGCCGGTGTCTC TGAGCAGCG GCCAGGCAGA GGCAGCTGTA CAAGGGCAGG
 481 GGCTTCTCTGC CGCTCTCTCA CCTCTGCCCC ATGCTGCCCG TGACCCCGGC AGAGACCGGG
 541 GACCTCAGGG ACCACTTGGAA GTCCGACATG TGCTCTCGC CCTGGAGAC CGACAGCATG
 601 GACCCGTTTG GGATGCCAC CAGACTTGGG GTGGTGAAGA GTCCCAGCTT TCAGAAATGA

Choloepus hoffmanni gene coding sequence (sloth) FGF19 (SEQ ID NO: 60) (Ensembl Accession No. ENSCHOT00000002324, which is hereby incorporated by reference in its entirety)

1 TTGCTCGAAA TGAAGGCAGT GGGCCTGCGG GCCGTGGCCA TCAAGGGCGT GCACAGTGCT
 61 CTGACCTCT CTGATGAACGC CGACCGCAGT CTGACCGGGC TGCTCTCGTA CTCTGAGAA
 121 GACTGTGCTT TTGAGGAGGA AATCCGCCCC GACGGCTACA ATGTTGACTG GTCTAGGAAG
 181 CACGGCCCTC CTGTCCTTT GAGCACTGCA AAACAGAGGC AGCTGTACAA AGGCAGAGGC
 241 TTTCTGCCCC TGTCCCACTT CCTGCCCCATG CTGCCCCATGA CGCCGGCCGA CGCCGAGAC
 301 CGCCGGGTAG ACCTGGAGTC GGACATGTTCT CTTTCACCTC TGAAACCGA CAGCATGGAT
 361 CCTTTGGAA TTGCTCTCCAG ACTTGAGCTT GTGAAACAGTC CAGCTTTAG CATAA

Ailuropoda melanoleuca gene coding sequence (giant panda) FGF19 (SEQ ID NO: 61) (GenBank Accession No. XM_002927906, which is hereby incorporated by reference in its entirety)

69 GG TCCTAGCGG CCTCTGCTG CGGGTAGCGG GGCGCCCCCT AGCCTTCTCG
 421 GACGGGGGG CGCACGTGCA CTACGGCTGG GGTGAGGCCA TCCGCTACG GCACCTGTAC
 481 ACCGCCGCCC CCCACGGCT CTCAGCTGC TTCTGCGCA TCCGTGCCGA CGGGGGGTT
 541 GACTGCGCG GGGGCCAGAG CGCGCACAGT TTGGTGGAGA TGAGGGCAGT CGCTCTGCG
 601 ACCGTGGCCA TCAAGGGTGT GCACAGCGC CGTACCTCT GCATGGCGC GGACGGCAGG
 661 ATGCAAGGGC TGCCCTAGTA CTCTGCGAGG GACTGTGCTT TGCGAGGAGGA GATCCGCCCC

TABLE 2-continued

721 GACGGCTACA ATGTGTACCG GTCCAAGAAG CACCGTCTCC CGCTCTCT GAGCGGTGCC
 781 AACAGAGGC AGCTTTACAA AGACAGAGGC TTCTGCCCC TGTCCTACTT CTTGCCATG
 841 CTGCCCGGA GCCCAGCAGA GCCCAGGGAC CTCCAGGACC ATGCGGAGTC GGACGGTTT
 901 TCTGCACCCC TAGAACAGA CAGCATGGAC CCTTTGGGA TCCGCACCAA AATGGGACTA
 961 GTGAAGAGTC CCAGCTTCCA GAAATAA

Sus scrofa gene coding sequence (pig) FGF19 (SEQ ID NO: 62) (Ensembl Accession No. ENSSCT00000014068, which is hereby incorporated by reference in its entirety)

1 ATCGGGAGCG CTCCGAGCC GTGGCGGGTG GTCCGCGGCC TGGTCTGGC CGGCCTCTGG
 61 CTGGCCGAG CGGGCGGCC CCTAGCCTTC TCGATGCTG GGCGCAGCT GCACACTCGC
 121 TGGGGCAGT CGGTCGGCT GCGCACCTG TACACTCGGA GTCCCCACGG CGTCTCCAGC
 181 TGCTTCTCTGC GCATCCACTC AGACGGCCCC GTGGACTGCTG CGCCGGGACA GAGCGCAGC
 241 AGTTGATGG AGATCAGGGC ATGCGGCTG AGTACCGTGG CGATCAAGGG CGAGCGCAGC
 301 GGCGTTACC TCTGCATGGG CGCCGACGGC AAGATGCAAG GGCAGACTCA GTACTCGGAT
 361 GAGGACTGTG CTTCGAGGA GGAGATCCCGC CCTGATGGT ACACAGTGTG CTGGTCCAAG
 421 AAACACCATC TGCCCGTGC TCTGAGCAGC GCGAGGAGA GGCAGCTGTA CAAGGGAGG
 481 GGCTTCTTGC CGCTGCTCCA CTTCCTGCC ATGCTGTCCA CTCTCCAGC CGAGCCGGAG
 541 GACCTCCAGG ACCCTTCAA GTCCGACCTG TTTCCTTGC CCCTGGAAAC GGACAGCATG
 601 GACCCCTTCC GGATCGCGC CAAACTGGGA GCGGTGAAGA GTCCAGCTT CTATAAATAA

Bos taurus gene coding sequence (bovine) FGF19 (SEQ ID NO: 63) (GenBank Accession No. XM_599739, which is hereby incorporated by reference in its entirety)

406 ATGCG GAGCGCTCCG
 421 AGCGGGTGC CGTGGCCCG CGCCCTGGTC CTGGCTGGCC TCTGGCTGGC CGCAGCGGG
 481 CGCCCCCTGG CCTTCCTCGA TGCGGGGCCG CACGTGCACT ACGGCTGGGG CGAGTCGGTT
 541 CGCTTGCAGC ACCTGTATAC CGCGGGCCCG CAGGGCCTCT ACAGCTGCTT TCTGCGCATC
 601 CACTCGACG CGCCGCGA CTGCGCGAG CTCAGAGCGG CGCACAGTTT GATGGAGATC
 661 AGGGCGGTG CGTCTGAGCAC CGTAGCCATC AAGGGCGAGC GAGCGTGCT GTACCTCTGC
 721 ATGACGCGG ACAGCAAGAT GCAAGGACTG ACCAGTACT CAGCCGAGGA CTGTGTTTC
 781 GAGGAGGAGA TCCGCTCTGA CGGCTACAAC GTGACTGTG CCAGGAAGCA CCATCTCCCG
 841 GTCTTCTCTG CGACGCTCCAG CGACAGGAGC CTGTTCAAAA GCAGGGGCTT CCTGCCGCTG
 901 TCTCACTTC TGCCCATGCT GTCCACCATC CCAGCCGAGC CTGAAGACCT CCAGGAACCC
 961 CTGAAGCTGTG ATTCTTTCTG GCCCCGTAAA ACAGATAGCA TGGACCTTT CGGGCTCGCC
 1021 ACCAAACTGG GATCGGTGAA GAGTCCCAGC TTCTATAATT AA

Canis lupus familiaris gene coding sequence (dog) FGF19 (SEQ ID NO: 64) (GenBank Accession No. XM_540802, which is hereby incorporated by reference in its entirety)

1 CTAGCCTTCT CGGACGGGG GCGCACGTC CACTCCTCT GGGGGGAGCC CATCCGCTG
 61 CGGCACCTGT ACACCGCCGG CCCAACGGC CTCTCCAGCT GCTTCTGCG CATCCGCGCC
 121 GACGGCGGG TGACTCGCG CGCGGGCCAG AGCGGCACA GTCTGATGGA GATGAGGGCG
 181 GTCGCTCTGC GGACCGTGGC CATCAAGGGC GTGCACAGCG GCGGTACCT CTGCATGGGC
 241 GCGGACGGCA GGATGCAAGG GCTGCTCTAG TACTCCGCG GAGACTGTAC TTTCGAGGAG
 301 GAGATCGCTC CGCATGGCTA CAATGTGTA TGCTCAAGA ACACCATCT CCCCATCTCT
 361 CTGAGTAGTG CCAAACAGAG CGACGCTCTAC AAGGGCAGGG GCTTTTTGCG CCTGTCAC
 421 TTCTTACCTA TCTTGGCCCG GAGCCAAACA GAGCCAGGG ACCTGGAAGA CCATGTGGAG
 481 TCTGACGGGT TTTCGCTAC CCTGGAAACCA GACAGCATGG ACCCTTTTGG GATCGCCACC
 541 AAAATTGGAC TAGTGAAGAG TCCCAGTTTC CAAAATAA

Oryctolagus cuniculus gene coding sequence (rabbit) FGF19 (SEQ ID NO: 65) (GenBank Accession No. XM_002724449, which is hereby incorporated by reference in its entirety)

1 ATCGGGCCGG CGCCGAGCGG AGGTGGCGCG GCCCGCGCCT TGGTCTGGC CGGCCTCTGG
 61 CTGGCCGCGG CGCGCGGCC CCTGGCTTG TCCGACGCGG GCGCGCATCT GCACACTCGC
 121 TGGGGCGAGC CGGTCGGCT GCGCACCTG TACGCCACCA GCGCCCGACGG CGTCTCGCAC
 181 TGCTTCTCTGC GTATACGCG CGACGGCGCC GTGGACTGCG AGCGGAGGCC GAGCGCACAC
 241 AGTTGCTGG AGATCCAGC CGGTCGGCTG CGCACCGTGG CCTTCAAGGG CGTGCACAGC
 301 TCCCGCTACC TCTGCATGGG CGCGACGGC AGGATGCGGG GCGAGCTGCA GTACTCGGAG
 361 GAGGACTGTG CCTCCAGGA GGAGATGAGC TCCGCTACA ACCTGTAACCA GACAGAGGG
 421 CACCATCGTCT CGGTGCTCT GAGCAGTGCC AAGCAGAGAC ACCTGTAACCA GACAGAGGG
 481 TTCTGCCCG CCTCCCAACT CTCGGCGTG CTGGCCCTGG CCTCGAGGA GACCGCGGCC
 541 CTCGGCGACC ACCCTGAAGC CGACCTGTT CCCCCGGCCC TGAAACCGA CAGCATGGAC
 601 CCCTTCGGCA TGGCCACCAA GCTCGGGCCG GTGAAGAGCC CCAGCTTTCA GAAGTAG

Pteropus vampyrus gene coding sequence (megabat) FGF19 (SEQ ID NO: 66) (Ensembl Accession No. ENSPVAT00000009907, which is hereby incorporated by reference in its entirety)

1 ATGCGGAGCC CGTGGCTGT GGCGCGGCC TTGGTCTGG CGGGCTCTG GCTGGCCTCA
 61 GCTGCGGGCC CCTCGCCCT CTGGACGCG GGGCGGCACG TGCACTACGG CTGGGGCGAG
 121 GCCATCCGCC TGCGGCACCT GTACACGCC GGGCCCCACG GCGCCCTCAG CTGCTTCTG
 181 CGCATCCGG CGGATGGGC GTGGACTGCG CGCGGGGCC AGAGCGGCCA CAGTTGGTG
 241 GAAATCCGGG CTGTCGCCCT GCGGAACAGTG GCTATCAAGG GCGTGCACAG CGTCCGATAC
 301 CTCTGCATGG GAGCCAGCG CAGGATGCTA GGCTGCTTC AGTACTCCGC TGACGACTGC
 361 GCCTTCGAGG AGGAGATCCG CCGGACGGC TACAACTGTG ACTACTCCAA GAAGCACAC
 421 CTCCCGTCT CTCTGAGCAG TGCCAGCAG AGGCAACTGT ACAAGGGACAG GGGCTTCTG
 481 CCCCTGTCCC ATTCTCTGCC CATGCTGCC AGGAGCCCGA CAGAGCCCGA GAACCTCGAA
 541 GACCACTTGG AGGCCACAC GTTTCTCG CCCCCGGAGA CAGACGACAT GGACCCCTTT
 601 GGGATTGCCA GTAAATTGGG GCTGGAGGAA AGTCCAGCT TCCAGAAAGTA A

TABLE 2-continued

Tursiops truncatus gene coding sequence (dolphin) FGF19 (SEQ ID NO: 67) (Ensembl Accession No. ENSTRT00000000066, which is hereby incorporated by reference in its entirety)

```

1 ATGCCGAGCG CTCCGAGCG GTGCCGCCTG GCCCGGCCGG TGTTCCCTGG CGGCCCTCTGG
61 CTGGCTCGAG CGGGCGCCC CCTAGCCTTC TCGGATGCCG GGCGGCACGT GCACTACGGC
121 TGGGGCGAGT CGGTCGGCCT GCGGCACCTG TACACCGCGG GTCCCCAGGG CCTCTCAGGC
181 TGCTTCCCTGC GCATCCACTC AGACGGCGCC GTGGACTCTGC CGCCGGTTCA GAGCGCGCAC
241 AGTTTGATGG AGATCAGGGC AGTGCCTCTC AGTACCGTGG CCATCAAGGG CGAACCGAGC
301 GTCTCTGACC TCTGCATGGG CGCCGACGGG AAAATGCAAG GCTTGAGTCA GTACTCAGCT
361 GAGGACTGTG CTTTGAGGA GGAAATCCGT CGGGACGGT ACACAGTGTAA CTGGTCCAAG
421 AAACACCACC TCCCCTGTC CCTGAGCAGC GCCAGGCAGC GGCAGCTGTT CAAAGGCAGG
481 GGTTTCTGTC CGCTGTCTCA CCTCTCTCCC ATGCTGTCCA CCATCCCCAC AGAGCCCGAT
541 GAAATCAGG ACCACTTGA GCGGATTTG TTGCTTGC CCCTGAAAAC AGATAGCATG
601 GACCCATTG GCCTCCGCAC CAAACTGGGA GTGGTGAAGA GTCCCAGCTT CTATAAGTAA

```

Myotis lucifugus gene coding sequence (microbat) FGF19 (SEQ ID NO: 68) (Ensembl Accession No. ENSMLUT00000002508, which is hereby incorporated by reference in its entirety)

```

1 ATGCAAAGCG CGTGGAGCGC ACGCGTTGTG GCCCGAGCCC TGGCTTGGC CAGCCTCGGG
61 CTGGCTCTAG CGGGGGGGCC CCTCGGTCTT TCGGACGCTG GGCGGCACGT GCACTACGGC
121 TGGGGGGAGT CCATCCCGCT GCACCGACCTG TACACCTCCG GCCCCCACGG CCCATCCAGC
181 TGCTTCCCTGC GCATCCCGCG TCACGGCGCA GTGGACTCTGC CGCGGGGCCA GAGCGCGCAC
241 AGTTGGTGG AGATCAGGGC CGTCGCCTG CGGAAAGTGG CCATCAAGGG CGTGCACAGC
301 GCCTGTACCT TCTGCATGGG AGGGACAGGGC AGGATGCTGG GGCTGCCCTCA GTTCTCGCCC
361 GAGGACTGTG CTTTGAGGA GGAGATCCCG CGGGACGGCT ACAACGTGTA CGGGTCCAG
421 AAGCCACAGC TGCCCGTCTC GCTGAGCAGT GCCCGCGAGA GGCAAGCTGTT CAAGGCCCGG
481 GGCTTCTGTC CGCTGTCTCA CCTCTGCCCC ATGCTGCCCA GCAGCCCCGC GGGACCCGTG
541 CCCGAGAGC GCCCTCTGGA GCGGACAGAG TTCTCTTCGC CCCTGGAAAAC AGACAGCATG
601 GACCCATTG GGATTGCAA CAACCTGAGG CTGGTGAGAA GTCCCAGCTT TCAGGAATAA

```

Ornithorhynchus anatinus gene coding sequence (platypus) FGF19 (SEQ ID NO: 69) (GenBank Accession No. XM_001506664, which is hereby incorporated by reference in its entirety)

```

1 ATGCTTCTCT GTGGGGTTT GCCTAGTCTG CTGGAGATCA AGGCGGTGGC CGTGCACAGC
61 GTGCCCATCA AAGGGGTCCA CATCTCTCGG TACCTCTGC TGAAGAGAGA TGGGAAACT
121 CCATGGGCAC GTCTGCTGGA GATCAAGGGCG GTGGCGTGC GCACGGTGGC CATCAAAGGG
181 GTCCACAGCT CTGCTGACCT CTGATGGAA GAGGATGGAA AACTCCATGG CGAGATTGG
241 TATTCTCGAG AAGACTGTG TTTGAAGAG GAAATACGTC CAGATGGCTA CAATGTGTAT
301 AAATCTAAGA AATATGGTGT TCCCTTTCT TTAAGCAGGG CCAAACAAAG GCAGCAATT
361 AAAGGAAGAG ACTTCTGC TCTTCTCGT TTCTTGCCAA TGATCAACAC AGTGCCTGT
421 GAGCCAGCAG AGTTGGGAA CTATGCCAT TACTTTGAAT CAGATATATT TTCCCTCACCT
481 CTGGAACAGC ACAGCATGG CCCATTAGA ATTGCCCTA AACTGTCCCC TGTAAGAGC
541 CCCAGCTTC AGAAATAA

```

Monodelphis domestica gene coding sequence (opossum) FGF19 (SEQ ID NO: 70) (GenBank Accession No. XM_001373653, which is hereby incorporated by reference in its entirety)

```

1 ATGGCCCAGC TCCCTGCCCC GCTCCTCACC CTGGCTGCTC TCTGGCTGGC CCCGACGGCG
61 CGTGCCTGGC CGCTGGTGGA CGCCGGGCCT CACGTCTACT ACGGCTGGGG GGAGGCCATT
121 CGTCTCGGGC ATCTCTACAC GGCAATCGG CACGGCGCTC CGAGCTTCTC CTTCCTCGG
181 ATCCCCCGGC AGGGCCCGT GGACGGCAGC CGGAGTCAGA GGCGCCTAG TTTGCTGGAG
241 ATCAAGGGCG TAGCTCTCG GATGGTGGCG ATCAAAGGT TCCATAGCTC TCGGTACTG
301 TGTTGGAG ACGGCGGGAA ACTCCAGGGA TCGGTGAGGT TCTCGGCCGA GGACTGCA
361 TTGAGGAGC AGATTCCTCC CGACGGCTAC AACGTGTACT AGTCCCCCAA GTACAACCTC
421 CCCGCTCTGC TCTGACTGA CAAGCAGAGG CAGCAGGCC ACGGCAAGGA GCACCTGCC
481 CTGCCCCACT TCCGCCAT GATCAATGCT ATTCTTTGG AGGCCGAGGA GCCCGAGGGC
541 CCCAGATGT TGGCGGCCCTC TCTGGAGACG GACAGCATGG ACCCCTTCGG CCTCACCTCC
601 AAGCTGTTGC CGGTCAGAG CCCAGCTT CAGAAATAA

```

Anolis carolinensis gene coding sequence (anole lizard) FGF19 (SEQ ID NO: 71) (GenBank Accession No. XM_003214667, which is hereby incorporated by reference in its entirety)

```

1 ATGTGTCGGC GGGCGTTGCC TCTGCTGGGG GCCCTCTGG GCTTGGCGGC CGTGGCTCC
61 CGCGCCCTCC CGCTCACCGA CGCCGGGCC CACGTCAAGT ACGGCTGGGG GGAGGCCGTC
121 CGGCTCAGGC ACCTCTACAC CGCGGGGGCGG CAGGGCCTCT TCAAGCCAGTT CCTCCGCATC
181 CACGCCGAGC GGAGAGTCGA CGGGCCCGGC AGCCAGAAC GGCAGAGTTT GCTGGAGATC
241 CGCGGGCTCT CGTGGCGCGC CGTGGCCCTC AAAGGGCTGC ACAGCTCCCC CTACCTCTGC
301 ATGGAGGAGG ACGGCGGCT CGCGGGGATG CTCAGATATT CTGAGAAGA CTGTTCTTT
361 GAAGAGGAGA TGCGTCCAGA TGGCTACAAT ATCTACAAGT CAAAGAAATA CGGAGTTTG
421 GTCTCTCTAA GTAATGCCAG ACAAAAGACAG CAACTCAAG GGAAAGATTT TCTTCCTTTG
481 TCTCATTTCT TGCGCATGAT CAAACATGTG CCAGTGGAGT CTGAGACTT TGGAGAGTAT
541 GGTGACACCA GGCAGCATTGA TGAATGGAT ATTTCAGTT CACGTCTGAA AACTGACAGC
601 ATGGACCCCTT TTGGCCTCAC TTCAGAAGTG TCATCAGTAC AAAGTCCTAG CTTTGGAAA
661 TAA

```

Ochotona princeps gene coding sequence (pika) FGF19 (SEQ ID NO: 72) (Ensembl Accession No. ENSOPRT00000010769, which is hereby incorporated by reference in its entirety) (1-214, excluding 78-112)

```

1 GTGCGGAGCA GGGGAGCCAT GGCCGCCTG CTGGTTCTAG CCACTCTCTG GCTGGCGCG
61 ACAGGGCGGC CGCTGGCCTT GTCCGACGCG GGGCGCACC TGCACTACGG CTGGGGCGAG

```

TABLE 2-continued

121 CCCATCCGCC TCGGGCACCT GTACGCCACC AGCGCCCACG GCCTCTCGCA CTGCTTTTG
 181 CGCATCCGTA CCGACGGCAC CGTGGACTGC GAGCGCAGCC AGAGCGGCCA CA-----

 242 -----CTAC AGTACTCGGA GGAGGACTG
 266 GCCTTCGAAG AGGAGATCG AGCTGGCTAT AACGTGTACC GCTCCAGGAG GTACCACTG
 326 CCCGTGTCCC TGGCAGCGC CAGGAGAGG CAGCTGCAGC GGAGCGGTGG CCTTCCTGCC
 386 CTGTCCCCACT TCCTGCCGGT GCTGCCGCG GCCTCGGAGG AGGTGGCGGC CCCCCTGAC
 446 CACCCGAAG CAGACCCCTT CTCGCCCTG GAGACCGACA GCATGGACCC ATTTGAAATG
 506 GCCACCAAGC GGGGGCTGGT GAAGAGCCCC AGCTTCCAGA AGTGA

Cavia porcellus gene coding sequence (guinea pig) FGF19 (SEQ ID NO: 73) (Ensembl Accession No. ENSPCOT00000008222, which is hereby incorporated by reference in its entirety)

1 ATATGGAGTG CGCCGACGG AGTGTGGTG ATCCCGGCC TGGTCTGGC TGGCTGTGG
 61 CTGGCGGTGG CGGGCGCCC CCTGGCCCG CGGCTCTCG CGCTATCTGA CCAGGGGCCG
 121 CACTTGACT ACAGCGTGGA CCAGCGATC CGCCCTCGGC ACCTGTACGC CGCGGGGCC
 181 TACGGCGCT CGCGCTGCTT CCTCGCATT CACACGGAGC GGCGGGTGGA CTGCGTCGAG
 241 GAACAGAGCG AGCACTGTTT GCTGGAGATC AGAGCAGTCG CTCTGGAGAC CGTGGCCATC
 301 AAGGACATAA ACAGCGTCG GTACCTGTGC ATGGGCCCG ACGGCAGGAT CGGGGCCCTG
 361 CCTCTGGTATT CGGAGGAGGA CTGTGCTTC AAGGAAGAGA TCAGCTACCC GGGCTACAGC
 421 GTGTACCGCT CCCAGAGCA CCACCTCCC ATCGTGTGA GCAGTGTCAA GCAGAGGCAG
 481 CAGTACCGAGA GCAAGGGGGT GGTGCCCCCTG TCCTACTTCC TGCCCATGCT GCCCAAGGCC
 541 TCTGTGGAGC CCAGCGACGA GGAGGAATCC AGCGTGTCT CGTTGCCCT GAAGACGGAC
 601 AGCATGGACC CCTTGGAT GGCCAGTGAG ATCGGGCTGG TGAAGAGTCC CAGCTTTCAG
 661 AAGTAA

Tupaia belangeri gene coding sequence (tree shrew) FGF19 (SEQ ID NO: 74) (from Ensembl Accession No. ENSTBET00000000307, which is hereby incorporated by reference in its entirety) (1-219, excluding 116-138)

1 ATAGAGGAGAA CACCGACGGG GTTGCAGTG GCCCGTGTCC TCTTCCTGGG CAGCCTTTGG
 61 CTGGCGCAG CGGGGAGCCC CCTGGCCCTG TCCGACGCCG GGCGCATGT GAACCTACGGC
 121 TGGGATGAGT CCATACGCC GCGCACATTG TACACCGCA GCCCGCACGG CTCCACCAGC
 181 TGCTTCTTGC GCATCCCGTCA CGACCGCTCA GTGGAACCTGGG CGGGGGGCCA GAGTTTGAC
 241 AGTTGCTGG AGATCAAGGC AGTCCCTTG CAGACCGTG CGATCAAAGG CGTGTACAGT
 301 GTCCGCTACC TCTGCATGGA CGCGACGGC AGGATGCAGG GGTG----- NNGGTCCACG
 361 AAGCACGGC TCCCAGTCTC CCTGAGCAGT GCCAAGCAGA GGCAAGCTGTT AACGGTTAGG
 429 GGCTTCCTT CCCTTCCTCA CCTTCTGCTC ATGATGGCCA AGACTTCAGC AGGGCCTGGA
 489 AACCCAGGG ACCACCCAGG GTCTAACACT TTCTGTTGC CCTTGGAAAC TGATAGCATG
 549 GACCCATTG GGATGACCAC CAGACATGGG CTGGTGAAGA GTCCAGCTT TCAAAACTAA

Rattus norvegicus gene coding sequence (Norway rat) FGF15 (SEQ ID NO: 75) (GenBank Accession No. NM_130753, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 56)

1 ATGGCGAGAA AGTGGAGTGG GCGCATTGTC GCGCGAGCTC TGGTCTGGC CACTCTGTGG
 61 CTGGCCGTGT CTGGCGTCC CCTGGTCCAC CAATCCCAGT CTGTGTCCGA TGAAGGTCCA
 121 CTCTTCTCT ATGGCTGGG CAAGATTACCC CGCCCTGCAGT ACCTGTACTC TGCTGGTCCC
 181 TACGCTCTCA ACTGCTTCTC GCGTATCCGG AGTGCAGGCT CTGTGGACTG CGAGGAGGAC
 241 CAGAACGAC GAAATCTGTT GGAGTCCCGC CGGTTGCTC TGAAGACAAT TGCCATCAAG
 301 GACGTCAGCA CGCTGGCGTA CCTCTGCATG AGCGCCGAC GCAAGATATA CGGGCTGATT
 361 CGCTACTCGG AGGAAGACTG TACCTTCAGG GAGGAATGG ACTGTTTGGG CTACAACCAAG
 421 TACAGGTCCA TGAAGCACCA CCTCCACATC ATCTTCATCA AGGCCAACGCC CAGAGAGCAG
 481 CTCAGGGCC AGAAAACCTTCA AAACATTATTC CCCATATTTC ACCGGTCTTT CTTTGAATCC
 541 ACGGACCCAGC TGAGGTCTAA ATATGTTCTC CTGCCCTGG AGAGCGACAG CATGGATCCG
 601 TTCAGAATGG TGGAGGATGT GGACCACCTA GTGAAGAGTC CCAGCTTCCA GAAATGA

Mus musculus gene coding sequence (house mouse) FGF15 (SEQ ID NO: 76) (GenBank Accession No. NM_008003, which is hereby incorporated by reference in its entirety)

148 ATG GCGAGAAAGT GGAACGGGGC TGCGGTGGCC
 181 CGAGCCTGG TCCCTGGCAC TCTGTGGCTG GCTGTGTCTG GGCGTCCCTT GGCTCACCAA
 241 TCCCAGTCTG TGTCAAGATGA AGATCCACTC TTCTCTACG GCTGGGGCAA GATTACCCGG
 301 CTGCGAGTACCG TGTACTCCGC TTGTCCTAT GTCTCCAAGT GCTTCCTCCG AATCCGGAGC
 361 GACGGCTCTG TGGACTCGCA GGAGGACCAA AACGAACGAA ATTGTTGGA ATTCCGGCG
 421 GTGCGCTCTGA AGACGATTGC CATCAAGGAC GTCAAGCAGG TGCGGTACCT CTGCATGAGC
 481 CGGGACGGCA AGATATACGAG GCTGATTGCG TACTCGGAGG AAGACTGTAC CTTCGGGAG
 541 GAAATGGACT GTTAAAGCTCA CAAACAGTAC AGATCCATGA AGCACCATCT CCATATCATC
 601 TTCACTCCAGG CCAAGCCAG AGAACAGCTC CAGGACCAGA AACCCCTAA CTTTATCCCG
 661 GTGTTCTACC GCTCTTCTT TGAACACGGG GACCAAGCTGA GTGCTAAAT GTTCTCCCTG
 721 CCCCTGGAGA GTGACAGCAT GGATCCGTT AGGATGGTGG AGGATGTAGA CCACCTAGTG
 781 AAGAGTCCCA GCTTCCAGAA ATGA

Gallus gallus gene coding sequence (chicken) FGF19 (SEQ ID NO: 77) (GenBank Accession No. NM_204674, which is hereby incorporated by reference in its entirety)

127 ATGG GGCGCCGCC CCCCGCGCA CCCGGCGCTG CCCTGGCGCT GCTGGGGATC
 181 GCCGCCGCC CGGCCGCCGC CAGGTCCTG CGCGTCCCG ACCTCGGGGG TCCGCACGTC
 241 AACTACGGCT GGGGGAAACC CATCCGGCTG CGGCACCTAC TACACCGGCC AGGCAAGCAC
 301 GGGCTCTCA GCTGCTTCTC GCGCATCGGC GGCGACGGCC GGGTGGACGC TGTCGGTAGC
 361 CAGAGCCGC AGAGTCTGTT GGAGATCCGC GCCGTGGCGG TGCGCACCGT GGCCATCAAG

TABLE 2-continued

421 GCGGTGCAGA GCTCCCGCTA CCTCTGCATG GACGAGGC GGCGGCTGCA CGGGCAGCTC
 481 AGCTATTCCA TTGAGGACTG TTCTCTTGAA GAGGAGATTG GTCCAGACGG CTACAACGTG
 541 TATAAATCAA AGAAATACGG GATATCGGT TCTTGAGCA GTGCCAAACAA AAGACAGCAA
 601 TTCAAAGGAA AGAGTTTCTC CCGCTGTCT CACTTCTTAC CCATGATCAA CACTGTGCCA
 661 GTGGAGGTGA CAGACTTTGG TGATATATGGT GATTACAGCC AGGCTTTGA GCCAGAGGTC
 721 TACTCATCGC CTCTCGAAC GGACACGATG GATCCCTTG GGATCACTTC CAAACTGTCT
 781 CCAGTGAAGA GCCCCAGCTT TCAGAAATGA

Taeniopygia guttata gene coding sequence (zebra finch) FGF19 (SEQ ID NO: 78) (GenBank Accession No. XM_002194457, which is hereby incorporated by reference in its entirety)

1 ATGGTTATCA TAAGCAATCT ATATCTGATG CAGAACGATG TTATGATGAA TATGAGGCAGA
 61 GCACCCCTC CGCTCACCGC TGCTCGCTC TCGGCCACCC CTGCGCTCCG GCTGCCGCTG
 121 CGCCGCCCG ACGCCGGCCC GCACCTCAA TACCGCTGGG GAGAGCCCAT CGGGCTGCCG
 181 CACCTCTACA CGGCCAGCAA GCACGGGCTC TTCACTGCTC TCCCTGCTAT CGGCCTGAC
 241 GGCGGGTGG ACGCCGGCGG CAGCAGAGC CCAGCAGAGCC TGCTAGAGAT CGGCCTGCTG
 301 CGCGTGCAGC CCGTGGCCAT CAAGGGCGTG CAGAGCTGGT GGTACCTGTG CATGGACGAG
 361 CGCGGGCGGC TGCACTGGCA GCTCAGGAAT TCCACTGAAG ACTGCTCTT TGAGGAGGAG
 421 ATTGCCCCAG ACGGTACAA TGTGTATAGA TCTAAAAAAC ATGGAATATC GGTGCTTTG
 481 AGCAGTGCCA AACAAAGACA GCAGGTCAAG GGGAAAGATT TCCCTTCCCT GTCTCACTTC
 541 TTGCCATGCA TCAACACTGT GCCCATGGAG TCAGCAGACT TTGGTGAATA TGTTGATTAC
 601 AGCCAGGCTT TTGAGGAGA GGCCTCTCC TCACCTCTGG AGACGGACAG CATGGACCCC
 661 TTGGCATCG CCTCCAAACT GTCCCTAGTG AAGAGCCCTA GCTTCCAAAAA CTGA

Danio rerio gene coding sequence (zebrafish) FGF19 (SEQ ID NO: 79) (GenBank Accession No. NM_001012246, which is hereby incorporated by reference in its entirety)

1 ATGCTCTCT TACTCTTTGT CACTGTTGT GGAAGTATCG GCGTGGAGAG CCTCCCGTTG
 61 CCCGACTCTG GTCCACATTT GGCAATGAC TGAGTGAAG CGCTCCGGCT ACGACATCTG
 121 TACCGAGCA GACATGGCTT ACATCTGCAA ATAACACACG ACCGAGAAAT CATTGGATCC
 181 ACATGCAAG CTCGGACAGT AAGTTTGTGAG GAGATATGCC CGGTGGACAC AGGCTCGTA
 241 GCCATTAAGG GAGTTGCAAG CTCCCGATTT CTTGATGG AAAGACTGGG AAACCTGTAC
 301 GGATCGACAA TTACACTAA AGGAGACTG TCTTTTTGGG AAGGCATCTT TCCAGACGGC
 361 TACAACGTC ACTTCTGAG CAAACACGGG GCTCTTGTA CTTTAAGTGG TGCGAAAAAC
 421 AAGTTGCACA GTAACGATGG GACTCTGCA TCCAGTTCC TCCCCATGAT CAACACACTT
 481 TCAGAGGAAC ACACTAAACA GCACTCAGGG GAACAGCACT CTTCTGTTAA CCATGGACAG
 541 GACCATCAGT TTGGCTCTGA AATAGACAGT ATGGACCCCTT TCGGAAAGAT CTCTCAAATA
 601 GTGATCCAGA GTCCAGCTT CAACAAAAGA TGA

Xenopus (Silurana) tropicalis gene coding sequence (Western clawed frog) FGF19 (SEQ ID NO: 80) (GenBank Accession No. NM_001142825, which is hereby incorporated by reference in its entirety)

1 ATGTTGAAAGA CCTGCCCTTG GATTGTTGGT CCCATGATGG TGGCCGTGCT GTATTTCTC
 61 GGAGGGCGG AAAGTCTGCC GCTTTTGAT GCCGGGCCGC ACATGCAGAA CGGCTGGGG
 121 GAGTCATGCA GAATTCGGCA CCTGATAGC GCGCAGGAGT TCGGGCAGGA CAGCTACTAC
 181 CTCCGGATAC ACAGGATGG CAGAGTCGAT CGTACAGGG AAAAACCAT CCACAGTTA
 241 TTGAAATCA GAGCAATTG AGTTGGAATT GTTGCCTTA AAGGGTATCG CAGCTCTTG
 301 TACCTGTGCA TGGGGTCCGA GGGAAACTC TATGGAATGC ACAGTTACTC CCAGGATGAT
 361 TGCTCTTTG AAGAGGAGT TCTCCGGAT GGATACAACA TGATATAATC AAGGAAACAT
 421 GGCCTGTGCT CTCCTCTGA CAAAGGAG CAGAGCAAC AATACAAAGG AAAGGGCTAC
 481 CTCCGGTGT CCCATTCT ACCCGTGATA AGCTGGGTGC CCATGGAGCC CACCGAGAT
 541 GTAGAAAGATG ATATCTACAG GTTCCATTC AATACGGACA CAAAAAGTGT CATTGACAGC
 601 CTTGATACCC TGGGACTAAT GGATTTTCG AGTTATCACA AGAAATAG

Otolemur garnettii (bushbaby) FGF19 gene coding sequence (SEQ ID NO: 81) (Ensembl accession no. ENSOGAT00000031686, which is hereby incorporated by reference in its entirety)

1 ATGCCAGCG GGCTGAGAGG GCGTGTGGTA GCGGGGCCCGG TGGCCCTGGC CAGCTCTGG
 61 CTGGCCGTGG CGGGGCCCGG GCTGGCCTTC TCAGGATGCCG GCCCTCACGT GCACACTGGC
 121 TGGGGTGAGC CCATCCGCTC GCGCACCTG TACACCGCCG GCCCCCACGG CCTCTCAGC
 181 TGCTCTCTGC CGGTACCCAC CGACGGTGCG GTAGACTGCG CGCGGGGCCA GAGCGCACAC
 241 AGTTTGTGCTT AAATCAGGGG CGTGTCTCGT CGGACCGTGG CCATCAAAGG CGTGCACAGC
 301 CGCCGGTACCGC TCTGCAATGGG CGCCGACGGC AGGATGCAAGG GGCTGCCCTCA GTACTCGAG
 361 GAAGACTGTG CCTTGAGGA GGAGATCCGG CCAGACGGCT ACAACGTCTA CTGGCTGTGAG
 421 AAGCACCGCC TGCCGGTGTCTC TCTGAGCTGAGT GCGCCGCAAGA GGAGCTGTA CAAGGGCAGG
 481 GGCTCTCTGC CGCTCTCTCA CCTCTCTGCG ATGCTGCTCG TGACCCCGAC CGAGCCCCGG
 541 GACCTCAGAG ACCACCTGGA ATCCGACATG TTCTCTTTGC CCTGGAAAC TGACAGCATG
 601 GATCCATTG GGATCGCAC CAGACTGGGC GTGGTGAAGA GTCCAGCTT TCAGAAATGA

Felis catus (cat) FGF19 gene coding sequence (SEQ ID NO: 82) (Ensembl accession no. ENSFCAT00000026317, which is hereby incorporated by reference in its entirety)

1 ATGGGGAGGG CGCCGAGCCA GTGGCGGGTA ACCCGCGCCCG TGCTCTCTAGC CGGTCTCTGG
 61 CTGGCAGCAG CGGGGCCCGG CCTAGCCTTC TCAGGACCGGG GGCCTCACGT GCACACTGGC
 121 TGGGGTGAGC CCATCCGCTC GCGCACCTG TACACCGCCG GCCCCCACGG CCTCTCAGC
 181 TGCTCTCTGC CGATCCGAGC CGACGGGGGG GTGACTGCG CGCGGAGCCA GAGCGCACAC
 241 AGTTTGTGAGT AGATCAGGGC AGTCGCTCTG CGGACCGTGG CCATCAAAGG CGTGCACAGC
 301 GTCCGGTACCGC TCTGCAATGGG CGCCGACGGC AGGATGCAAGG GGCTGCCCTCA GTACTCTGCT
 361 GGGGACTGTG CCTTCAAGA GGAGATCCGC CCCGACGGCT ACAATGTGTA CGGGTCCGAG
 421 AACGACCGTC TCCCGTCTC TTGAGTAGT GCCATACAGA GGAGCTGTA CAAGGGCAGA

TABLE 2-continued

481 GGTTTTGCG CCTGTCCCCTTCTTGCAGCAGCAGG
 541 GACCTCCAGG ACCACGTGGAGGGAGGTTCTTACACCTGGAAACAGACAGCATG
 601 GACCCTTTGGATGGATTGCCACCAAATGGGGTTAGTGAAGA GTCCCAGCTT CCAAAAGTAA

Pelodiscus sinensis (Chinese softshell turtle) FGF19 gene coding sequence (SEQ ID NO: 83) (Ensembl accession no. ENSPISIT00000010427, which is hereby incorporated by reference in its entirety)

241 ATGTGGAG GAGCCTGTGC AAATCTCACAA
 301 CGTCTCTGGC TCTGCTGGGAGCTCTGTTTCGGTGTGAGATCTCG CTTTCTCGG
 361 ATGCAGGGCC ACATGTGAAC TATGGCTGGGGAGGACCTATCGATTAAAGG CACCTATACA
 421 CCGCCAGCAG ACACGGGCTG TTCAATTACT TCTGAGGAT CAGCAGTGTGGCAAAGTGG
 481 ATGGCACAGCAGTTCAGAGT CTCACAGTC TGCTGGAAAT CAGGGCTGTGCGAGTGG
 541 CGGTGGCGAT CAAGCTTCCC GGTAACCTCTG CATGGAAGAA GACGGGAAGC
 601 TGCATGGACT TCTCAGGTAT TCTACAGAACAGTGGCTTGAAGAGGAG ATACGCCAG
 661 ATGGCTACAAATGTTATAAACTCAAGAAAT ATGGAATCTCTGTTAAGTGCCTAAGTAGTGC
 721 AACAAAGACA ACAATTCAAAGGAAAGACT TTCTTCCATT GTCTCACTTC TTGCTATGA
 781 TCAATACAGT ACCTGTGGAG TCAATGGATT TTGGAGAATA TGTTGATTAT AGTCATACTT
 841 TTGAAATCAGA TCTATTCTCTC ACCTCTCG AAACGTGACAG CATGGATCCC TTGGAATCA
 901 CCTCTAAAT ATCTCCAGTG AAGAGCCCCA GCTTTCAGAAATAA

Latimeria chalumnae (coelacanth) FGF19 gene coding sequence (SEQ ID NO: 84) (Ensembl accession no. ENSLACT00000014697, which is hereby incorporated by reference in its entirety)

1 ATGTTACAGG CACTGTACAA TCTCTGTACA CCTCTAGTTT TGTTTAAGCT TCCCTTTGCA
 61 ATGGGGGGTACACCCCTGCC TTCTGCCAAAT GAAGGGCCCC ATCTGAACCTA TGACTGGGA
 121 GAATCTGTAA GACTCAAACA TCTGTACACA TCTAGCAAGC ATGGATTGAT CAGTTACTT
 181 TTACAGATCA ATGATGATGG CAAAGTAGAT GGGACCACTA CACGAAGCTG TTATAGTTG
 241 CTCGAAATAA AATCAGTGGG GCCAGGAGTT TTGCAATTAA AAGGCATACAGAGCTTCCAGA
 301 TACCTTGTG TCGAGAAGGA TGGAATATTG CATGGATCGGC GCACATTATTC AGCAGACGAT
 361 TGCTCTTCA AAGAGGATAT ACTCCAGAT GTTACACTA TCTACGTGTC AAAGAAACAT
 421 GGATCTGTGTTAATCTGAG CAACACAAAA CAGAACAGTC AGAGAAATCG CAGAACCTG
 481 CCTCCATTTCCTCAGTTCTC ACCGCTTATG GACACCATTG GTGTGGAGTG CATGAACACTGC
 541 GGGGAGCACT GTGACGACAA CCTGCATGAC GAGCTAGAAA CAGGACTGTC CATGGATCCC
 601 TTGAAAGTA CATCAGAAAA ATCCTTTCAG AGTCCCAGCT TTCACAATAG ATAA

Mustela putorius furo (ferret) FGF19 gene coding sequence (SEQ ID NO: 85) (Ensembl accession no. ENSMPUT00000004650, which is hereby incorporated by reference in its entirety)

421 ATGCCGG AGCGCCCGGA GTCGCTGCCGCGTGGACCCCG GCGCTGGTCC TAGCCGGCT
 481 TTGGCTGGCC GCAGCCGGGC GCCCCCTAGC CTCTCGGAC GCGGGGGCCGACGTGACTA
 541 TGGCTGGGGT GAGCCCATCC GCCTACGGCA CCTGTACACC GCGGGCCCC ACGGCCTCTC
 601 CAGCTGCTTC CTGCGCATCC GTGGCAGCG CGGGGTTGAC TGCGCGCGG GCCAGAGCGC
 661 GCACAGTTG GTGGAGATCC GGGCAGTGC TCTGGGAGC GTGGCCATCAAGGGCTGTA
 721 CAGCGACCGC TATCTCTGCA TGGGTGCGGA CGGCAAGGATG CAGGGCTGCTCCTAGTACTC
 781 CGCGGAGAC TGTGCTTTCG AGGAGGAGAT CGGCCCTGTGAGGTACAACG TGACGGTC
 841 CAAGAAGCAC CGTCTCCCG TCTCCCTGAG CAGTGCAGAA CAAAGGGAGC TGACAGGAGA
 901 CCGGGGCTTT TTGCTCTGT CCCATTCTT GCCCATGCTG CCCGGGAGCC TGGCGGAGCC
 961 CAGGGACCTC CAGGACACAG TGGAGGCTGA TGTTTTCTG GCCCCCTAG AAACAGACAG
 1021 CATGGACCTT TTGGAATG CCACCAAAAGGGACTAGTG AAGAGTCCCA GCTTCCAAAA
 1081 ATGA

Takifugu rubripes (fugu) FGF19 gene coding sequence (SEQ ID NO: 86) (Ensembl accession no. ENSTRUT00000007155, which is hereby incorporated by reference in its entirety)

1 TCATCTACAA GGATTAAGTGG AACACATGGCTT CTCCCTCATGC TCCCCATCAC CGTTGCAAAC
 61 CTCTTCTCT GTGCTGGAGT TCTCTCTTG CCTTGTGTTGG ATCAAGGGTC TCATTTTCCC
 121 CAAGGCTGGG AACAGGTAGT CGCGCTTCAGG CACCTGTATG CTGCGAGTGC AGGGCTGCAC
 181 CTGCTGATCA CTGAAGAGGG CTCGATCCAA GGCTCTGAG ATCCAACCTTT ATACAGCCTG
 241 ATGGAGATCC GTCCGGTGGAGCC CAGGAGCTGT GTGTCATTAAGGAGCAGC AACCAACAGC
 301 TTCCCTGCA TAGAAGGTGCTGGAGACTG TACTCATCAC AGACCTACAG CAAAGACGAC
 361 TGACCTCTCA GAGAGAAAT CCTAGCAGAC GGCACAGCGG TCTACAGATC TGCGACAC
 421 GGAGCTCTGG TCAGTCTGGG AAAACTACCGG CAGCAGCTGA GGGGGAGGA CTGGAGCGTT
 481 CCGACACTGG CTCAGTCTCTC CCCCAGAATA AGTTCACTGG ATCAGGACTT TAAAGCTGCT
 541 CTGACGAGA CTGAGAAGGCC AGAACAAACT GCACCTCAAA GATCGGAACCTGTCAGATG
 601 GTGGACTCAT TTGGAAGACT CTCTCAGATC ATCCACAGTC CCAGTTTCA CAAG

Equus caballus (horse) FGF19 gene coding sequence (SEQ ID NO: 87) (Ensembl accession no. ENSECAT00000021494, which is hereby incorporated by reference in its entirety) (1-216, excluding 1-19 and 114-216)

1 -----GCG
 4 GCCGGCGCC CCCTAGCCTT GTCCGACGCT GGGCCGACAGC TGCACTACGG CTGGGGCGAG
 64 CCGATCCGCC TCGGGCACCT GTACACCGCC GGGCCCCACG GCCTCTCCAG CTGCTTCTG
 124 CGCATCCCGCG CCGATGGCGC CGTGGACTGCG GCGCGGGGCC AGAGCGCGCA CAGTTGGTGC
 184 GAGATCAGAG CAGTCGCTCT GCGCACCGTG GCCATCAAGG GCGTGCACAG CGTCCGGTAC
 244 CTCTGCATGG CGCAGCGCAGG CAGGATGCAA GGGCTGGTA

TABLE 2-continued

Oryzias latipes (medaka) FGF19 gene coding sequence (SEQ ID NO: 88)
 (Ensembl accession no. ENSORLT00000000352, which is hereby incorporated by reference in its entirety)

```

1 ACCATGCTGC TCATGTGGT CACCATTTCC ACAATGGGT TTTCTGACTC TGAGATTTC
61 AGCATGCCG TCTCTGATCA TGGACCCCAC ATCACTACA GCTGGAGCCA AGTGGTCCGC
121 CTCCGGCAC C TGTACGGGT CAAGCCTGGC CAACATGTC AGATCAGAGA GGATGGACAC
181 ATCCACGGCT CAGCAGAAC AACTCTGAAC AGCCTGCTGG AGATCCGTC GGTTGCTCCG
241 GGACGGTGG TCTTCAGAGG AGTAGCCACC TCAAGGTTTC TGTCATGGA GAGCGACGGC
301 AGACTCTTC CTCACACAC ATTGACAAG GACAACCTGGC TCTTCAGAGA GCAGATCTTG
361 GCAGACGGCT ACAACATCTA CATTTCAGAT CAGCATGAA CCTCTGCTTAG TTTGGGAAAC
421 CACCGGAAA GGCAGCAGGG TTAGACCGG GATGTTCCAG CCCTGGCTCA GTTCCTCCCC
481 AGGATCAGCA CCCTGAGCA GGGCGTGTAC CCAGTGCAG ACACCCCCCA CCAGATGAGA
541 ACAATGCAA CAGAGAAC TCTAGATGCC ACAGACACAT TTGGCAACT CTCTAAAATC
601 ATTACACAGTC CCAGCTTCA CAAAAGATGA
  
```

Xiphophorus maculatus (platyfish) FGF19 gene coding sequence (SEQ ID NO: 89) (Ensembl accession no. ENSXMAT00000001519, which is hereby incorporated by reference in its entirety)

```

1 ATG
4 TTTGTGTTCA TTCTATGCGAT TGCTGGTGA CTTTTTACTC TGGGAGTATT TTGCATGCCA
64 ATGATGGACCA AGGGGCCAT TGTCACCCAT GGATGGGGCC AGGTGGTCCG GCACCGGCAT
124 CTGTATGAG CCAAGCAGG ACTGCACCTA CTGATCAGTG AGGTGGACAA AATCCACGGT
184 TCCGCAGATC AAACCTTTA CAGCCTGCTG GAGATCCAACT CTGTTGGCCC CGGACGTGTT
244 GTGATCAAG GAGTGGCAC CACACGCTTC CTCTGCATGG AGAGCGACGG CAGATTGTAC
304 TCAACTGAAA CATAACAGCAG AGCTGACTGC ACCTTCAGAG AACAGATCCA CGCAGACGGC
364 TACAACGTT ACACCTGTA TAGGCATGGA GCCTCTCTCA GTTTGGGAAAC CAACCGACAA
424 AGACACAGCG GCTCAGACGG TTGTTTCCA GCTCTGGCCC GCTTCTTCC CAGGTTAAAC
484 ACCCTTCAGC AGGCCGCTCC CACAGAGCCG GATGTTCTG ATCAGCTCAG TCCAGAGAAA
544 GTACAACAGA CTGTTGGACAT GTGTCCTCC TTGTCAG TCTCTCATAT AATTACAGT
604 CCCAGCTTCC ATAAGAGATG A
  
```

Ictidomys tridecemlineatus (squirrel) FGF19 gene coding sequence (SEQ ID NO: 90) (Ensembl accession no. ENSSTOT00000026298, which is hereby incorporated by reference in its entirety)

```

1 ATGCGGAGCG CGCCGAGCGG ACCTGCTTAA GCCTGGGCC TGGTGGCTGGC CAGCCTCTGG
61 TTGGCAGTGG CGGAGCACC CCTGGGCCGG CGCTCTCTGG CTCTCTCCGA CCAGGGGCCA
121 CACTTGACT ATGGCTGGGA TCAGGCCATC CGCCTCCGGC ACCTGTACGC CGCGGGGCC
181 TACGGCTTCT CCAACTGTTT CCTGCGCATC CGCACCGACG GCGCCGTGGA CTGCGAGGAG
241 AAGCAGAGCG AGCCTGTTT GATGGAGATC AGGGCGGTGG CTCTGGAGAC TGTGGCCATC
301 AAGGACATAA ACAGCGTCCG TAATCTCTGC ATGGGCGCCG ACGGCAGGAT ACAGGGACTG
361 CCTCGGACT CGGAGGAAGA GTGCAGTTTC AAGGAGGAGA TCAGCTATGA CGGCTACAAAC
421 GTGTACCGGT CCCAGAGTA CCACCTTCCC GTGTTGCTCA GCAGTGCACAA CGAGCGGCAG
481 CTGTACCAGA GCAAGGGCGT GGTTCCTCTG TCTACTTCC TGCCCATGCT GCCCCTGGCC
541 TCTGCGGAGA CGAGGACCG CTTGAATCC GATGTTCTCT CTTTACCTCT GGAAACTGAC
601 AGCATGGACC CGTTGGGAT GGCGAGTGAA GTGGGCCTGA AGAGCCCCAG CTTCCAGAAG
661 TAA
  
```

Gasterosteus aculeatus (stickleback) FGF19 gene coding sequence (SEQ ID NO: 91) (Ensembl accession no. ENSGACT00000018770, which is hereby incorporated by reference in its entirety)

```

1 ATGCTGCTGC TGCTGGTCCC CGCGTACGTT GCCAGTGTGT TTTAGCTCT CGGGGTTGTT
61 TGCTTGCCTTAAACAGATCA GGGCTCCAC ATGGCCGAGC ACTGGGGCCA GTCGGTCCGA
121 CTCAAGCACC TGTACGCCG CAGCCGGGA CTCCACCTGC TGATGGGGGA GGATGGTGG
181 ATCCAAGCT CCGCGCAGCA AAGCCCTAC AGCCTGCTGG AGATCAGTGC AGTGGATCCG
241 GGCTGTGTTG TCATCAGAG AGTAGCAACCG GACAGGTTTC TCTGCATCGA AGGCAGATGGA
301 AGACTGACT CATGGACAC STACACAGAGA GACGACTGCA CTTTCAGGGGA GCAGATCTC
361 CGGACGGCT ACAGCGTCTA CGTCTCCCAT GGACACGGGG CCTCTGCTCAG CCTGGGGAAC
421 CACAGGCAGA GGCTGCGAGG TCGAGACCCAC GGCCTGGCC CGCCACCGC CGCCACCGAG
481 AGGGTCAGCA CCATGGATCA GGCTCTGGCC CGCCACGGCC CGCCACCGAG CGCCACCGAG
541 ACGGAAGAGC CGCTGGACTC GTTGGAAAG CTCTCTCAGA TCATTCAAG TCCCAGCTTC
601 CACGAGAGAT GA
  
```

Oreochromis niloticus (tilapia) FGF19 gene coding sequence (SEQ ID NO: 92) (Ensembl accession no. ENSONIT00000022816, which is hereby incorporated by reference in its entirety)

```

55 ATGCTG
61 CTGCTCCTCA TCGTATCCAT TGCAATATG CTTTTGGTG TTGGAATGGT TTGCATGCC
121 CTGTCAGACA ACGGGCCCA CATCGCCAC GGTGGGCC AGGTGGTCCG GTCAGGCAC
181 CTTTACCGCA CCAGACGGG AATGCACTG CTGATCAGTG AGGGTGGACA GATCCGTGGT
241 TCTGCGTCC AGACTCTGCA CAGCTTAATG GAGATTGTC CAGTCGGTCC AGGCCGTGTT
301 GTCATCAGAG GGGTAGCAAC CGCAAGGTTT CTCTGCATAG AAGACGACGG CACACTGTAC
361 TCTACCGACG CCTACAGCAG AGAGGACTGC ATCTTCAGAG AGCAGATCTT CGCAGATGGG
421 TACAACATCT ACATCTCTGA CAGACATGGA GTCTGCTCA GTCTGGGAAAC CAACCGGCAA
481 AGACTGAGC GCTTAGACCG AGGAGATCCA GCCTGGGCC AGTCCCTCCC CAGGATCAGC
541 ACTCTGAATC AAATCCCTC CCCTGGGCA AACATCGGTG ACCACATGAA AGTAGCAAAA
601 ACAGAAGAAC CTGTTGGACAC AATAGATTCA TTGGAAGT TCTCTCAGAT CATTGACAGT
607 CCCAGCTTCC ATAAGAGATG A
  
```

TABLE 2-continued

Meleagris gallopavo (turkey) FGF19 gene coding sequence (SEQ ID NO: 93) (Ensembl accession no. ENSMGAT00000011114, which is hereby incorporated by reference in its entirety) (1-216, excluding 1-70)

```

1 GTAGGCAATC AATCACCCACA GAGCATCCTT GAAATAACTG CTGTTGATGT CGGGATCGTC
61 GCTATCAAGG CTTGTTCTC TGGCAGATAC CTGGCCATGA ACAAAAGGGG CAGGTTTAT
121 GCATCACTCA GCTATTCCAT TGAGGACTGT TCCTTGAAAG AGGAGATTCG TCCAGATGGC
181 TATAACGTGT ATAAATACAA GAAATACGGG ATATCAGTGT CTTTGAGCAG TGCAAACAA
241 AGACAACAAT TCAAAGGAAA AGATTTCTC CCACTGTCTC ACTTCTTAC CATGATCAAC
301 ACTGTGCCAG TGGAGGTGAC AGACTTGGT GAATACGGTG ATTACAGCCA GGCTTTGAG
361 CCAGAGGTCT ACTCATGCC TCTCGAACG GACAGCATGG ATCCCTTGG GATCACTTCC
421 AAACGTCTC CAGTGAAGAG CCCAGCTT CAGAAA

```

Papio anubis (olive baboon) FGF19 gene coding sequence (SEQ ID NO: 94) (GenBank accession no. XM_003909422, which is hereby incorporated by reference in its entirety)

```

758 ATG AGGAGCGGGT GTGTTGGTGGT
781 CCACGCCCTGG ATCCTGGCCA GCCTCTGGCT GGGCGTGGCC GGGCGTCCCC TCGCCTTCTC
841 GGACGCCGGG CCCCACGTGC ACTACGGCTG GGGGACCCC ATCCGCCTGC GGCACCTGTA
901 CACCTCCGGC CCCCACGGGC TCTCCAGCTG CTTCCTGCAC ATCCGCACCG ACGGCGTCGT
961 GGACTGCGGG CGGGGCCAA GCGGCCACAG TTGCTGGAG ATCAAGGCAG TAGCTCTGCG
1021 GACCGTGGCC ATCAAGGCC TGACAGCGT GCGGTACACTC TGCATGGCG CCGACGGCAA
1081 GATGCAAGGGG CTGCTCTAGT ACTCAGAGGA AGACTGTGT TTCGAGGAGG AGATCCGCC
1141 TGATGGCTAC AATGTATACC GATCCCCAGA GCAACGCCCTC CCGGTCTCC TGAGCACTGC
1201 CAAACAGCGG CAGCTGTACA AGAACAGAGG CTTCTTCCG CTGCTCTCATT TCCTGCCAT
1261 GCTGCCCTATG GCCCCAGAGG AGCCTGAGGA CCTCAGGGGC CCCTTGGAAAT CTGACATGTT
1321 CTCTTCGCCC CTGGAGACTG ACAGCATGGA CCCATTGGG CTTGTCAACG GACTGGAGGC
1381 GGTGAGGAGT CCCAGCTT CAGAAATAA

```

Saimiri boliviensis boliviensis (Bolivian squirrel monkey) FGF19 gene coding sequence (SEQ ID NO: 95) (GenBank accession no. XM_003941165, which is hereby incorporated by reference in its entirety)

```

231 ATGCGGAGCG
241 GGTGTGTGGT GGTCCACGCC TGGATCCTGG CTGGCCTCTG GCTGGCTGTG GTGGGGGGCC
301 CCCTCGCCCTT CTCCGATGCG GGGCGCATG TGCAATTACGG CTGGGGCGAC CCCATTGCGC
361 TGCGGCACCT GTACACCTCC AGCCCCCACG GCCTCTCCAG CTGCTTCTG CGCATCGCA
421 GCGACGGCGT CGTGGACTGC GCGGGGGGCC AGAGCGCGCA CAGTTTGCTG GAGATCAAGG
481 CAGTCGCTCT AAGGACCGTG GCCATCAAGG GCGTGCACAG CTCGCGGTAC CTCTGCATGG
541 CGGCCGACGG CAGGCTGCAG GGGCTGTTCC AGTACTCGGA GGAAGACTGT GCTTTCGAGG
601 AGGAGATCCG CCCCGACGGC TACAATGTGT ACCTATCCGA GAAGCACCGC CTCCCGGTCT
661 CCCTGAGCAG CGCCAAACAG CGGCAGCTGT ACAAGAAACG AGGCTTTCTT CCGCTGTCCC
721 ATTCTCTGCC CATGCTGCC AGAGCCCCAG AGGAGCTGAA TGACCTCAGG GGCCACTTGG
781 AATCTGACGT GTTCTCTCA CCCTGGAGA CTGATAGCAT GGACCCATTG GGGCTTGTCA
841 CGGGACTCGGA GCGGGTGAAC AGTCCCCAGCT TTGAGAAGTA A

```

Pteropus alecto (black flying fox) FGF19 gene coding sequence (SEQ ID NO: 96) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org))

```

1 ATGCGCAGCC CGTGCAGCGGT GGGCGCGCG CTGGTGTGG CGGGCCTGTG GCTGGCGAGC
61 GCGGGGGGCC CGCTGGCGCT GAGCGATGCG GGGCCGCATG TGCAATTATGG CTGGGGCGAA
121 GCGATTGCGC TGCGGCATCT GTATACCGCG GGGCCGCATG GCGCGAGCAG CTGCTTCTG
181 CGCATTGCG CGGATGCGC GGTGATTG TGCGGGCGCC AGAGCGCGCA TAGCCTGTTG
241 GAAATTGCG CGGGCGCGCT GCGAACCGTG GCGATTAAGG CGGTGCATAG CGTGCCTAT
301 CTGTCATGG CGGGGATGG CGCGCATGCTG GGGCTGCTCG AGTATAGCGG GGATGATTG
361 GCGTTGAAAG AAGAAATTG CCGGGATGGC TATAACGTGT ATCATAGCAA AAAACATCAT
421 CTGCGGTGA GCCTGAGCAG CGCGAACACAG CGCCAGCTGT ATAAAGATCG CGGCTTCTG
481 CGCGTGCAGC ATTTCCTGCC GATGCTGCC CGCAGCGCGA CGAACCCGA AAACTTTGAA
541 GATCATCTGG AAGCGGATAC CTTAGCAGC CGCGTGGAAA CGCATGATAT GGATCCGTTT
601 GGCATTGCGA GCAAACCTGGG CCTGGAAAGAA AGCCCGAGCT TTCAAGAAA

```

Myotis davidii (David's myotis) FGF19 gene coding sequence (SEQ ID NO: 97) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org))

```

1 ATGAGCGGCC AGAACAGCGG CGCCCATGGC AGCCGCCCGG CGCTGGATGA AGAACCGGAA
61 CGGGGCCCGG TGGAACTGCG CGCGCTGGGC AGCACCGCG CGGATCCGCA CGTGTGCGAT
121 TTTCTGGAAA ACCATTTCT GGGCTATACC TGCCCTGAAAC TGGATATTG CCTGGCGACC
181 TATCTGGCG TGAGGACATG GGGGAAAGC ATTCCGCTGC GCCATCTGTA TACCAAGGGC
241 CGCGCATGCC CGAGCGAGCTG CTTCTGCGC ATTGCGCTGG ATGGCGCGGT GGATTGCGCG
301 CGCGGCCAGA CGCCGCATAG CCTGGTGGAA ATTCCGCGG TGCGCTGCG CAAAGTGGCG
361 ATTAAAGCGC TGCACTGCG GCTGTATCTG TGCACTGGAG CGCATGGCGG CATGCGGGC
421 CTGCCGCAGT TTAGCCCGA AGATTGCGCG TTGAGAAGAA ATTCGCCCC GGATGGCTAT
481 AACCTGTTAC GCAGCCAGAA ACATCAGCTG CGCGTGGAGCC TGAGCAGCGC CGGCCAGCGC
541 CAGCTGTTA AAGCGGCCGG CTTCTGCCG CTGAGCCATT TTCTGCCAT GCTGCCAGGC
601 AGCCGCCGG AACCGGTGCA TCGCGAACGC CGCGTGGAAAC CGGATGCGTT TAGCAGCCCG
661 CTGGAAACCG ATAGCATGGA TCCGTTGGC ATTGCGAACAA ACCTGCGCCT GGTGAAAAGC
721 CCGAGCTTTC AGAAA

```

TABLE 2 -continued

Tupaia chinensis (Chinese tree shrew) FGF19 gene coding sequence (SEQ ID NO: 98) (generated using SMS Reverse Translate tool on the ExPASY Bioinformatics Resource website (www.expasy.org/)) (1-257, excluding 13-19)

```

1 ATGCGCCGCA CCTGGAGCGG CTTTGCCTGT GCGAC----- CGCGCG
61 GGCAGCCCGC TGGCGCTGGC GGATCGGGC CCGCATGTGA ACTATGGCTG GGATGAAAGC
121 ATTGCGCTGC GCCATCTGTA TACCGCGAGC CTGCGATGGCA GCACCAGCTG CTTTCTGCGC
181 ATTGCGCATG ATGGCAGCGT GGGCTGCGCG CGCGGCCAGA GCATGCGATAG CCTGCTGGAA
241 ATTAAGCGG TGCGCTGCAGA GACCGTGGCC ATTAAAGCGG TGATATAGCGT CGCGTATCTG
301 TGCGATGGATA CGGATGGCGC CATGCGAGGC CTGCGCGACT ATAGCGAAAGA AGATTGACCC
361 TTTGAAGAAG AAATTGCGAG CGATGGCCAT AACGTGTATC GCAGCAAAA ACATGGCCTG
421 CGCGTGAGCC TGAGCAGCGC GAAACAGCGC CAGCTGTATA AAGGCCGGG CTTTCTGAGC
481 CTGAGGCAATT TTCTGCTGAT GATGCGAAAC ACCAGCGCGG GCCCGGGCAA CGCGCGCAT
541 CGCGCAACC CGCGCATCA CGCGCATCCG AACACCTTTA GCGTGCCTG CGAACACCGAT
601 AGCATGGATC CGTTGGCAT GACCACCCGC CATGGCCTGC TGCTGGATAG CTGCTGCGC
661 AGCCTGGTGC TGCTGACAT TAGGACCGAT GGGGAATTAA GCGCGTATGG CAACATTCTG
721 CGCCCGAGCT TTCGCTTTAA ACTGTTTAAATGAAAAAG TGACCAAC

```

Heterocephalus glaber (naked mole-rat) FGF19 gene coding sequence (SEQ ID NO: 99) (generated using SMS Reverse Translate tool on the ExPASY Bioinformatics Resource website (www.expasy.org/))

```

1 ATGCGCTTTA GCAAAAGCAC CTGCGCTTTT TTAAACCATC AGCGCCTGCAG GGCCTGTGG
61 CTGAGGCTGA GCAGCGTGA ATGGGTGCTG GATGCGCGG TGGAAAGGCCG CCCGATTGCG
121 CTGCGCCATC TGATGCGGC GGGCCCGTAT GGGCGCAGCC GCTGCTTTCTG CGGCATTCTAT
181 ACCGATGGCG CGGGTGGATTG CGTGGAAAGAA CAGAGCGAAC ATTCGCTTGCT CGAAATTGCG
241 GCGGGGGCGC TGGAAACCGT GGGCGATTAA GATATAACA CGCTGCGCTA TCTGTGCGAT
301 GGGCGGGATG CGCGCATGCA GGGCGTGGCG TGGTATAGCG AAGAAGATTG CGCGTTAA
361 GAAGAAATTA GCTATCGGG CTATAGCGTG TATCGCAGCC AGAACACATCA TCTGCCGATT
421 GTGCTGAGCA CGCTGAAAC GCGCCAGCGAG TATCAGAGCA AAGGCCTGGT GCGCTGAGC
481 TATTTCTGC CGATGCTGCC GAAAGCGAGC GTGGAACCGG CGCATGAAAGA AGAAAGCGC
541 TTAGCCTGC CGCTGAAAC CGATAGCATG GATCCGTTTG GCATGGCGAG CGAAATTGGC
601 CTGGCGAAAA GCGCGAGCTT TCAGAAA

```

Another member of the FGF19 subfamily, FGF21, is expressed primarily by the pancreas (Fon Tacer et al., "Research Resource: Comprehensive Expression Atlas of the Fibroblast Growth Factor System in Adult Mouse," *Mol Endocrinol* 24(10):2050-2063 (2010), which is hereby incorporated by reference in its entirety) and has metabolic effects similar to that of FGF19, such as increased energy metabolism, weight loss, lowered blood glucose levels, and resistance to obesity and diabetes (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6), 1627-1635 (2005); Coskun et al., "Fibroblast growth factor 21 corrects obesity in mice," *Endocrinology* 149(12): 6018-6027 (2008), which are hereby incorporated by reference in their entirety). Transgenic mice overexpressing FGF21 are also resistant to diet-induced obesity (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6), 1627-1635 (2005), which is hereby incorporated by reference in its entirety). Moreover, in diabetic rodent models, FGF21 administration lowers blood glucose and triglyceride levels (Kharitonenkov et al., "FGF-21 as a Novel Metabolic Regulator," *J Clin Invest* 115(6), 1627-1635 (2005), which is hereby incorporated by reference in its entirety).

In one embodiment of the present invention, the FGF21 portion of the chimeric protein of the present invention is from human FGF21 protein having an amino acid sequence of SEQ ID NO: 100 (GenBank Accession No. NP_061986, which is hereby incorporated by reference in its entirety) or a portion thereof, as follows:

30
In one embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein of the present invention comprises an amino acid sequence spanning residues corresponding to residues from position 29 to 167 of SEQ ID NO: 100, from position 29 to 190 of SEQ ID NO: 100, or from position 29 to 197 of SEQ ID NO: 100.

40
In one embodiment of the present invention, the N-terminal portion of the chimeric protein according to the present invention is or is derived from a mammalian FGF21.
45 In one embodiment of the present invention, the N-terminal portion of the chimeric protein according to the present invention is or is derived from a vertebrate FGF21. In one embodiment, the N-terminal portion of the chimeric protein according to the present invention is derived from a non-human vertebrate FGF21. It will be understood that this includes orthologs of human FGF21, or a polypeptide or protein obtained from one species that is the functional counterpart of a polypeptide or protein from a different species. In one embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein according to the present invention is derived from human, *pongo abelii*, *pan troglodytes*, *canis lupus familiaris*, *bos taurus*, *equus caballus*, *ailuropoda melanoleuca*, *oryctolagus cuniculus*, *gorilla gorilla*, *nomascus leucogenys*, *procavia capensis*, *cavia porcellus*, *tupaia belangeri*, *sorex araneus*, *ictidomys tridecemlineatus*, *loxodontia africana*, *sus scrofa*, *felis catus*, *otolemur garnettii*, *rattus norvegicus*, *mus mus-*

(SEQ ID NO: 100)

1 MDSDETGFEEH SGLWVSVLAG LLLGACQAH P IPDSSPLLQF GGQVRQRYLY TDDAQQTTEAH

61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA

121 CSFRELLLED GYNVYQSEAH GPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI

181 LAPQPPDVGS SDPLSMVGPS QGRSPSYAS

culus, vicugna pacos, anolis carolinensis, gadus morhua, latimeria chalumnae, tursiops truncatus, mustela putorius furo, takifugu rubripes, dipodomys ordii, echinops telfairi, macaca mulatta, microcebus murinus, ochetona princeps, xiphophorus maculatus, gasterosteus aculeatus, sarcophilus harrisii, macropus eugenii, xenopus tropicalis, danio rerio, bos grunniens mutus, saimiri boliviensis boliviensis, callithrix jacchus, tupaia chinensis, papio anubis, pteropus alecto, heterocephalus glaber, cricetus griseus, ovies aries, pan paniscus, macaca fascicularis, mesocricetus auratus, or oreochromis niloticus.

In one embodiment of the present invention, the portion of FGF21 of the chimeric protein of the present invention is from an ortholog of human FGF21 having an amino acid sequence as shown in Table 3. The portions of an ortholog of human FGF21 of a chimeric protein according to the present invention include portions corresponding to the above-identified amino acid sequences of human FGF21. Corresponding portions may be determined by, for example, sequence analysis and structural analysis. The high degree of FGF21 sequence conservation among mammals is shown in FIG. 9.

TABLE 3

<i>Pongo abelii</i> (Sumatran orangutan) FGF21 (GenBank Accession No. XP_002829565, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 101)	1 MDSDETGFEH SGLWVPVLAG LLLGACQAH IPDSSPLLQF GGQVRQRYLY TDDAQQTEAH
61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA	
121 CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPAPPEPPGI	
181 LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	
<i>Pan troglodytes</i> (chimpanzee) FGF21 (GenBank Accession No. XP_524333, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 102)	1 MDSDETGFEH SGLWVSVLAG LLLGACQAH IPDSSPLLQF GGQVRQRYLY TDDAQQTEAH
61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA	
121 CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPAPPEPPGI	
181 LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	
<i>Canis lupus familiaris</i> (dog) FGF21 (GenBank Accession No. XP_541510, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 103)	1 MGWAEAGFEH LGWVPVLAV LLLEACRAHP IPDSSPLLQF GGQVRQRYLY TDDAQETEAH
61 LEIRADGTVV GAARQSPESL LELKALKPGV IQILGVKTSR FLCQGPDGTL YGSLHFDPEA	
121 CSFRELLLED GYNVYQSETL GLPLRLRPHN SAYRDLAPRG PARFLPLPGL LPAPPPEPPGI	
181 LAPEPPDVGS SDPLSMVGPS QGRSPSYAS	
<i>Bos taurus</i> (bovine) FGF21 (GenBank Accession No. XP_001789639, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 104)	1 MGWDEAKFKH LGWVPVLAV LLGTCRAHP IPDSSPLLQF GGQVRQRYLY TDDAQETEAH
61 LEIRADGTVV GAARQSPESL LELKALKPGV IQILGVKTSR FLCQGPDGKL YGSLHFDPKA	
121 CSFRELLLED GYNVYQSETL GLPLRLPPQR SSNRDPAPRG PARFLPLPGL PAAPPDPPGI	
181 LAPEPPDVGS SDPLSMVGPS YGRSPSYAS	
<i>Equus caballus</i> (horse) FGF21 (GenBank Accession No. XP_001489202, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 105)	1 MDWDTGFKY QGLWVPVLAV LLLGACQSHP IPDSSPLLQF GGQVRQRHLY TDDAQETEAH
61 LEIRADGTVA GAVHRSPEL LELKALKPGV IQILGVKTSR FLCQGPDGTL YGSLHFDPEA	
121 CSFRELLLED GYNVYQSETL GLPLRLPHS SPYQDPAPRG PARFLPLPGL PPAPPEPPGI	
181 PAPEPPDVGS SDPLSMVGPS RSRSPSYAS	
<i>Ailuropoda melanoleuca</i> (giant panda) FGF21 (GenBank Accession No. XP_002917910, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 106)	1 MGWDEARSEQ LGLWVPVLAV LLLEACQAH IPDSSPLLQF GGQVRQRYLY TDDAQETEAH
61 LAIRADGTVV GAARSPEL LELKALKPGV IQILGVKTSR FLCQGPDGTL YGSVRDFPVA	
121 CSFRELLLED GYNVYQSETL GLPLRLPAHN SPYRDSAPRG PARFLPLPGL LPVPPDPPGI	
181 LGPEPPDVGS SDPLSMVGPS QGRSPSYAS	

TABLE 3-continued

Oryctolagus cuniculus (rabbit) FGF21 (Ensembl Accession No. XP_002723745, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 107)

1 MDWGKAKCRP PGLWVPALAA LLLGACQAH IPDSSPLLQF GDQVRQQHLY TDDAQTEAH

61 LEIRADGTVV GAARRSPESL LQMKAQPGI IQILGVQTSR FLCQRPDGTL YGSLHFDREA

121 CSFRELLRED GYNVYLSEAL GLPLRLSPGS SPRRAPAPRG PARFLPLPGL PPDLPEPPGGL

181 LAAAPPDVDS SDPLSMVQPA LDQSPSYTS

Gorilla gorilla (gorilla) FGF21 (Ensembl Accession No. ENSGGOP00000001229, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 108)

1 MDSDETGFEE SGLWVSVLAG LLLGACQAH IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA

121 CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPAPPEPPGI

181 LAPQPPDVGS SDPLSMVGPS QGRSPSYAS

Nomascus leucogenys (Northern white-cheeked gibbon) FGF21 (Ensembl Accession No. ESNLEP00000005639, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 109)

1 MDSDETGFEE SGLWVSVLAG LLLGACQAH IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA

121 CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPAPPEPPGI

181 LAPQPPDVGS SDPLSMVGPS QGRSPSYAS

Procavia capensis (hyrax) FGF21 (Ensembl Accession No. ENSOGAG00000001210, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 110)

1 MDWAKFGIEH PGLWVPVMAV LLLGACQGYP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIRADGTvv GAAHRSPEsl LELKALKPGI IQILGVKTSR FLCQGPDGVL YGSLRFDPVA

121 CSFRELLLED GYNVYQSEAH GLPLRLPSHN SPQRDLASRV PARFLPLPGR LTVLPEPSGV

181 LGPEPPDVGS SDPLSMVGPS QGRSPSYAS

Cavia porcellus (guinea pig) FGF21 (Ensembl Accession No. ENSCPOP00000000237, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 111)

1 MDWARTECER PRLWVSMLAI LLVGACQAH IPDSSPLLQF GGQVRQRYLY TDDAQDTEVH

61 LEIRADGSVR GIAHRSPEsl LELKALKPGV IQILGIRTSR FLCQRPDGSL YGSLHFDPEA

121 CSFRELLLED GYNVYKSEAH GLPLHLLRGD SLSQEPAAPG PARFLPLPGL PATPPEPPRM

181 LPPGPPDVGS SDPLSMVGPL WDRSPSYTS

Tupaia belangeri (tree shrew) FGF21 (Ensembl Accession No. ENSTBEP00000013946, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 112)

1 MGWDKARFEH LGAWAPVLA V LLLGACQAYP IPDSSPLLQF GGQVRQRYLY TDDTQDTEAH

61 LEIRADGTvv GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA

121 CSFRELLLED GYNIYQSEAR GLPLRLPPHD SPHRDRTPRG PARFLPLPGL PLVPPELPGV

181 LALEPPDVGS SDPLSMMGPS QQQSPSYAS

Sorex araneus (shrew) FGF21 (Ensembl Accession No. ENSSARP00000002784, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 113)

1 MVWDKARGQQ LGLWAPMLLG LLLGACQAH LPDSSPLLQF GGQVRLRFLY TDDAQRTGAH

61 LEIRADGTvQ GAAHRTPECL LELKALKPGV IQILGVTSR FLCQRPDGVL YGSLRFDPVA

121 CSFRELLLD GYNVYQSEAL GLPLYLHPPS APVSQEPASR GAVRFLPLPG LPPASLEPPR

181 PPAPVPPDVG SSDPLSMVGP PERHSPSYTS

TABLE 3-continued

Ictidomys tridecemlineatus (squirrel) FGF21 (SEQ ID NO: 114)
 1 MDWVKAKLEP LGLWVLVLAA LVLGACQAYP IPDSSPLLQF GGQVRQRYLY TDDAQETEAH

61 LEIRADGTVV GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQRPDGVL YGSLHFDPEA

121 CSFREQLLED GYNVYQSESH GLPVRLPPNS PYRDPAPPGP ARFLPLPGLP PAALEPPGIL

181 GPEPPDVGS DPLSMVGPLQ GRSPSYAS

Loxodonta africana (elephant) FGF21 (Ensembl Accession No. ENSLAFP00000016854, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 115)

1 MDWAKFGLE HPGLWVPVMA VLLLGACQGH PIPDSSPLLQF FGGQVRQRYLY YTDDQETEAH

60 LEIRADGTVA GAAHRSSESL LELKALKPGI IQILGVKTSR FLCQGPDGVL YGSLHFDPAA

120 CSFRELLLED GYNVYWSEAH GLPIRLPSHN SPYRDPASRV PARFLPLPGLP LPMLQEPPGV

180 LAPEPPDVDS SDPLSMVGPS QGRSPSYAS

Sus scrofa (pig) FGF21 (GenBank Accession No. NP_001156882, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 116)

1 MGWAEAKFER LGLWVPVLA VLLGACQARPI PDSSPLLQFG GQVRQRYLYT DDAQETEAHL

61 EIRADGTAG VARQSPESLL ELKALKPGVI QILGVQTSRF LCQGPDGRLY GSLHFDPEAC

121 SFRELLLEDG YNVYQSEALG LPLRLPPHRS SNRDLAPRGP ARFLPLPGLP PAPPEPPGIL

181 APEPPDVGS DPLSMVGPSH GRSPSYAS

Felis catus (cat) FGF21 (Ensembl Accession No. ENSFCAP00000006832, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 117)

1 MDWDEAGSQ RLGLWVVLGV LLPEACQAHF IPDSSPLLQF GGQVRQRFLY TDDAQETEVH

60 LEIKADGTVV GTARRSPESL LELKALKPGV IQILGVKTSR FLCQGPDGTL YGSLRFDPAA

120 CSFRELLLED GYNIYHSETL GLPLRLPPHN SPYRDLAPRA PARFLPLPGLP LPAPPEPPGI

180 LAPEPPDVGS SDPLSMVGPS QGRSPSYAS

Otolemur garnettii (bushbaby) FGF21 (Ensembl Accession No. ENSOGAG00000003581, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 118)

1 DKARTGFKH PGPWFPLLAV LLLGACQAHF IPDSSPLLQF GGQVRQRYLY TDDAQETEAH

60 LEIREDTGVV GAAQQSPESL LELKALKPGV IQILGVKTSR FLCQRPDGGL YGSLYFDPKA

120 CSFRELLLED GYNVYWSETY GLPLHLPPAN SPYWGPSLRS PARFLPLPGLP PAASPELPGI

180 LALEPPDVGS SDPLSMVGPS QGRSPSYAS

Rattus norvegicus (Norway rat) FGF21 (GenBank Accession No. NP_570108, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 119)

1 MDWMKSRVGA PGLWVCLLPL VFLLGVCEAY PISDSSPLLQ FGGQVRQRYLY YTDDDDQDTEA

61 HLEIREDTV VGTAHRSPES LLELKALKPG VIQILGVKAS RFLCQQPDTG LYGSPHFDPE

121 ACSFRELLLK DGYNVYQSEA HGLPLRLPK DSQDPATRGP VRFLPMPGLP HEPQEOPGV

181 PPEPPDVGS DPLSMVEPLQ GRSPSYAS

Mus musculus (house mouse) FGF21 (GenBank Accession No. NP_064397, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 120)

1 MEWMRSRVGT LGLWVRLLLA VFLLGVYQAY PIPDSSPLLQ FGGQVRQRYLY YTDDDDQDTEA

61 HLEIREDTV VGAHRSPES LLELKALKPG VIQILGVKAS RFLCQQPDTG LYGSPHFDPE

121 ACSFRELLLE DGYNVYQSEA HGLPLRLPK DSPNQDATSW GPVRFLPMPG LLHEPQDQAG

181 FLPPEPPDVG SSDPLSMVEP LQGRSPSYAS

TABLE 3-continued

Vicugna pacos (alpaca) FGF21 (Ensembl Accession No. ENSVPAP00000005562, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 121); partial sequence corresponding to human FGF21 residues 1 to 78, 169 to 171, and 183 to 209

1 MDWDEAKFEH RGLWVPVLTV LLLGACQARP IPDSSPLLQF GGQVRQRYLY TDDAQETEAH

61 LEIRADGTVV GVARQPE-----

121 -----GI P-----

181 --PEPPDVGS SDPLSMVGPS YSRSPSYTS

Anolis carolinensis (anole lizard) FGF21 (Ensembl Accession No. ENSACAP00000016895, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 122)

1 CKSKGGKGK ERMWVDLVFW AALLRTAPAL PLRNSNPIYQ FDGQVRLRHL YTADEQTHLH

61 LEILPDGTVG GSRPQNPFSL MEIKAVKPGV IRMQAKKTSR FLCMKPNGLR YGSLFYSEEA

121 CNFHEKVLS D GYNLYYSEN Y NIPVSLSSAG NLGQSRQLPP FSQFLPLVNK IPLEPVLEDF

181 DFYGHQLDVE SADPLSILGQ NPGFMSPSYV F

Gadus morhua (cod) FGF21 (Ensembl Accession No. ENSGMOP00000013789, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 123)

1 LLLLATLLHIG LSFYVPDSPGP LLWLGDQVRE RHLYTAESHR RGLFLEMSPD GQVTGSAAQT

61 PLSVLELRSV RAGDTVIRAR LSSLYLCVDR AGHTGQRQY TESDCTFREV ILEDGYTHFL

121 SVHHGLPISL APRHSPGRQG LRFSRFLPLR SSLSEDRVVAE PPDSPLNLDs EDPLGMGLGS

181 LLSPAFSM

Latimeria chalumnae (coelacanth) FGF21 (Ensembl Accession No. ENSLACP0000003781, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 124)

1 MLCQSFVILS QKFIFGLFLT GLGLTGLAWT RPFQDSNPIL QYSDSIRLRH LYTAESRHL

61 HLQINSQGV GGTTKQSPYS LLEMKA VKTG FVVI RGKKSA RYLCMERSGR LYGSLQYTEK

121 DCTFKEVVLA DGYNLYVSEE HQATVTLSPM RARIAQGKKI PPFSHFLPMV NKVPVEDVAA

181 EMEFVQVLRE MTADVDSPDP FGMTWEESVH SPSFFA

Tursiops truncatus (dolphin) FGF21 (Ensembl Accession No. ENSTTRP00000013808, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 125)

1 MGWDKTKLEH LGLWVPVLAV LLGPCQAHPI PDSSPLLQFG GQVRQRYLYT DDAQETEAHL

61 EIRADGTVVG TARRSPEGVK TSRFLCQGPE GRLYGS LHFN PQACSFRELL LEDGYNVYQS

121 EALGIPLRLP PHRSSNWDLA PRGPARFLPL PGFLPPP LEP PGILAPEPPN VGSSDPLSMV

181 GPSHGRSPSY TS

Mustela putorius furo (ferret) FGF21 (Ensembl Accession No. ENSMPUP00000003687, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 126)

1 MGWEARSEH LGLWVPVLAV LLLGACQAYP IPDSSPLLQF GGQVRQRYLY TDDAQETEAH

61 LEIRADGTVV GAARRSPESL LELKALKPGV IQILGVKTSR FLCQGPNGTL YGSFHFDPV

121 CSFREVLL ED GYNIYHSETL GLPLRLPPHN SPHRDLAPRG PARFLPLPGL LPATPESRG

181 PAPEPPNVGS SDPLSMVGPL QQQSPSYTS

Takifugu rubripes (fugu) FGF21 (Ensembl Accession No. ENSTRUP00000033950, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 127)

1 FIYLFIQTAL FSPSKWFNFY LPDSNPLSF DSHGRGIHLY TDNQRRGMYL QMSTDGSVSG

61 SDVQTANSVL ELKSVRN GH VIRGKSSLF LCMDSRGR LW GQRHPT EADC TFREVLLADG

121 YTRFLSLHNG TPVSLAPKQS PDQHTVBFTR FLPLRNTLAE ESMSEPPSNQ QRYFNIDSDD

181 LLGMDLNAMV SPQFSGDK

TABLE 3-continued

Dipodomys ordii (kangaroo rat) FGF21 (Ensembl Accession No. ENSDORP00000001155, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 128)
 1 MDQAKTRVGA RGLGGLVLAV IILGACKARP IPDSSPLLQF GGQVRLRHLY TDDTQETEAH
 61 LEIRADGTVV GTAHRSPESL LELKALKPGV IQILGIKTSR FLCQRPDGTL YGSLHFDPEV
 121 CSFQELLLED GYNIIYRSEAL GLPLRLSPDP APWGPARFLP LPGVPPAPPE PPGILAPEPP
 181 DVGSSDPLSM VGLLQGRSPS YAS

Echinops telfairi (lesser hedgehog tenrec) FGF21 (Ensembl Accession No. ENSETEP00000008707, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 129)
 1 MGCTKSGWKS PGLWVPVLAS LLGGCCGAHP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH
 61 LEIRADGTVG GVAHQSPKEF LSQWREKPLR SLHFDPAACS FREKLLEDGY NLYHSETHGL
 121 PLRLPREGD PSSQPGARFP PLPGQLPQLQ ETPGVLAPEP PDVGSSDPLS MVGPWRGQSP
 181 SYAS

Macaca mulatta (rhesus monkey) FGF21 (Ensembl Accession No. ENSMMUP00000031540, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 130)
 1 MDSDETGFEH SGLWVPVLAG LLGGACQAHP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH
 61 LEIREDTVG GAAHQSPES CGPEPGSEGG GAVGGAEGPG LLGLREAGLG PGSWLHFDPE
 121 ACSFRELLLE NGYNVYQSEA HGLPLHLPGN KSPHRDPASQ GPARFLPLPG LPPAPPEPPG
 181 ILAPQPPDVG SSDPLSMVGP SQARSPSYAS

Microcebus murinus (mouse lemur) FGF21 (Ensembl Accession No. ENSMICP00000012089, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 131)
 1 MGWDEAGAGF EHPGLWFPMI GVLLGACQA YPIPDSPLL QFGGQVRQRH LYTDDIQETE
 61 AHLEIRADGT VVGAARSPE LELKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEC
 121 SFRELLLEDG YNVYCPYLPL HLSPRIELAG SRSAALPPLA PERRILAPEP PDGSSDPLSM
 181 VGPSQGRSPS YAS

Ochotona princeps (pika) FGF21 (Ensembl Accession No. ENSOPRP00000006754, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 132)
 1 KDMMDGLQPPC LRVPVLAALL LGVGQARPIP DSSPLLQFGG QVRQRHLYTD DAQESEVHLE
 61 IRADGTAVGT ARRSPESSLE MKALKPGVIQ ILGVHTSRFL CQRPDGTLYG SLHFDHKACS
 121 FREQLLEDGY NVYHSETHGL PLRLSPDRAP RGPARFLPLP GPPPDLVPP LPPDVLAPEP
 181 DVVDSPDPLS MVGPLQGQSP SYTS

Xiphophorus maculatus (platyfish) FGF21 (Ensembl Accession No. ENSXMAP00000001576, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 133)
 1 CPFPFLFLIL SLPFFSSSFY IPESNPIFAF RNQLREHVLY TENHRRGLYV EIHLGDRVTC
 61 SDAQSPYSVL QIKSVKPGHV VIKGQTSSLF LCMDDSGNLR GQTTYDEADC SFRELLLADG
 121 YTRFLNSQHG VPLSLASRNS PDRHSVPFTR FLPLRNLTIV SEESTKTQRD FNLDSDLLG
 181 MG

Gasterosteus aculeatus (stickleback) FGF21 (Ensembl Accession No. ENSGACP00000010703, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 134)
 1 SLLLMLVPLPF CSSFYLTDS PLLPFNNQVK EVHLYTAENH RRAMYLQIAL DGSVSGSDAR
 61 STYSVQLQKS IQPGHVVIRG KASSMFLCVD SGGRRLRGQGP YSEADCSFRE LLLGDGYTRF
 121 LSSQHGSPLS LASRPSPDPN SVPFTRPLPI RTAPEAESVI EEPPSNQRYV NVDSEDLLGM
 181 GLNTVVSPQF SA

TABLE 3-continued

Sarcophilus harrisii (tasmanian devil) FGF21 (Ensembl Accession No. ENSSHAP00000005963, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 135); partial sequence corresponding to human FGF21 residues 3 to 172

1 VSAMGLRERA PRYLAPLLSSL LLACRASGHP LPDSSPMLLF GGQVRLRHLY TDVGQEAEAH

61 VELASDGTVR AAARRSPNSL LELKAVKPGI VRILAVHSSR FLCMRPNSEL YGAIHYDPSA

121 CNFRERLLGD GYNVYESEAH GRTLRLPPKA APGPAGPSRF LPLPG

Macropus eugenii (wallaby) FGF21 (Ensembl Accession No. ENSMEUPO0000013936, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 136)

1 TEEPSTGSRH LGQWAPGLPG PLLSLLLAYR GWGSPIPDSS PMLLFGGQVR LRHLYTDDQ

61 DTEAHVELGP DGVVRAVAER SPNSLLELKA VKPGVIRILA VQSSRFLCMR PNGELYGAVH

121 YDPSACNFRE HLLGDGYNRY ESETHRTTLR LSPSLGQAGP SRFLPLPGDW LPGPDPPWAQ

181 GPEPPDVGSADP DPLSMVGAVQ GLSPSYSS

Xenopus tropicalis (Western clawed frog) FGF21 (Ensembl Accession No. ENSXETP00000009917, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 137); partial sequence corresponding to human FGF21 residues 1 to 169

1 RGGRTKKKTL LRKWLCLLAI MLSRSRFSLA NPIQNSNPIL SNDNQVRTQY LYTDNNNMHL

61 YLQITHNGVV TGTEEKNDYG VLEIKAVKAG VVVIKGIRSN LYLCMDSRHQ LYASAYDKDD

121 CHFHEKITPD NYNMYSSEKH SEYVSLAPLK GSQMARFLPI

Danio rerio (zebrafish) FGF21 (Ensembl Accession No. ENSDARP00000094287, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 138)

1 MLLACFFIFF ALFPHLRWCY YVPAQNVLQ FGTQVRERLL YTDGLFLEMN PDGSVKGSPE

61 KNLNVCVLELR SVKAGETVIQ SAATSLYLCV DDQDKLGQH HYSALDCTFQ ELLLDGYSFF

121 LSPHTNLPVS LLSKRQKHGN PLSRFLPVSR AEDSRTQEVK QYIQDINLDS DDPLGMGHRS

181 HLQTVFSPSL HTKK

Bos grunniens mutus (yak) FGF21 (GenBank Accession No. ELR56628, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 139)

1 MGWDEAKFKH LGLWVPVLA VLLGTCRAHP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIRADGTVV GAARQSPESL LELKALKPGV IQILGVKTSR FLCQGPDGKL YGSLHFDPKA

121 CSFRELLLED GYNVYQSETL GLPLRLPPQR SSNRDPAPRG PARFLPLPGL PAEPPDPPGI

181 LAPEPPDVGS SDPLSMVGPS YGRSPSYTS

Saimiri boliviensis boliviensis (Bolivian squirrel monkey) FGF21 (GenBank Accession No. XP_003940375, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 140)

1 MGSEEVALER PALWVSVLAG LLLGTCQAYP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIREDTGTA GAAHQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLYFDPEA

121 CSFRELLLED GYNVYQSVAH SLPLHLPGR SPPWDPAPRG PARFLPLPGL PPEPPEAPGI

181 LAPEPPDVGS SDPLSMVGPS QQQSPSYTS

Callithrix jacchus (white-tufted-ear marmoset) FGF21 (GenBank Accession No. XP_003735669, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 141)

1 MGSEEVGLEH PALWVSVLAG LLLGTCQAH IPDSSPLLQF GGQVRQRYLY TDDAQKEAH

61 LEIXEDGTVA GAATKPKVS LLQLKALKPGV VIQILGVKTS RFLCQRPDGA LYGSLHFDP

121 ACSFRELLLE DGYNVYQSVH HGLPLHLPES RSPPRDAPR GPARFLPLPG LPPEPPEPPG

181 ILAPEPPDVGS SSDPLSMVGP SQQQSPSYAS

TABLE 3-continued

Tupaia chinensis (Chinese tree shrew) FGF21 (GenBank Accession No. ELW47159, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 142)

1 MGWDKARFEH LGAWAPVLAG LLLGACQAYP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIRADGTVV GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQRPDGAL YGSFHDPPEA

121 CSFRELLLED GYNVIYQSEAR GLPLRLPPHD SPHRDRTPQG PARFLPLPGL PLVPPELPGV

181 LALEPPDVGS SDPLSMMGPS QQQSPSYAS

Papio anubis (olive baboon) FGF21 (GenBank Accession No. XP_003915900, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 143)

1 MDSDETGFEH SGLWVPVLAG LLLGACQAHIP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIREDTGTVG GAAHQSPESK CGPEPGSEGG GALHFDPEAC SFRELLLENG YNVYQSEAHG

121 LPLHLPGNKS PHRDPAASGP ARFLPLPGLP PAPPEPPGIL APQPPDVGS DPLSMVGPSQ

181 ARSPSYAS

Pteropus alecto (black flying fox) FGF21 (GenBank Accession No. ELK18566, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 144)

1 MGWGKARLQH PGLWVPVLAG LLGACQAHPI LDSSPLQFG SQVRRRYLYT DDAQDTEAH

61 EIRADGTAVG AARRSPESLL ELKALKPGVI QVLGVKTSRF LCQRPDGTL YGSFHDPAA

121 SFRELLLDKG YNVYQSEALA RPLRLPPYSS PSSDPARRGP ARFLPLPGLP PEPPQPPGRL

181 APEPPDVGS DPLSMVWPSR GRSPSYTS

Heterocephalus glaber (naked mole-rat) FGF21 (GenBank Accession No. EHB06286, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 145)

1 MDWARAESER PGLWVPAVLA VLLLGACQAH PIPDSSPLLQ FGGQVRQRHL YTDDAQTEV

61 HLEIRADGSV GGAHRSPES LLELKALKPG VIQILGVRTS RFLCQRPDGTL YGSFHDP

121 ACSFRELLLA DGYNVIYQSEA YGLPLRMLPS DSASRDPVPP GPARFLPLPG LHPPPLEPPG

181 MLPPEPPDVGS SSDPLSMVGP LQGRSPSYAF

Cricetulus griseus (Chinese hamster) FGF21 (GenBank Accession No. XP_003508726, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 146)

1 MDWMKSGVGV PGLWVPPLLPI FLLGVQSQAH IPDSSPLLQF GGQVRHRHL TDDNQETEVH

61 LEIRQDGTVI GTTHRSPESL LELKALKPEV IPVILGVKASR FLCQQPDGTL YGSFHDPPEA

121 CSFRELLLED GYNVIYQSEVH GLPLRLPQRD SPNQAPASWG PVPPLPVPGL LHQPQELPGF

181 LAPEPPDVGS SDPLSMVGPL QGRSPSYAS

Ovis aries (sheep) FGF21 (GenBank Accession No. XP_004015845, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 147)

1 MGWDEAKFKH LGLWVPVLAG LLLGTCRAHP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIRADGTVV GAARQSPESL LELKALKPGV IQIFGVKTSR FLCQGPDGKL YGSFHDPKA

121 CSFRELLLED GYNVIYQSETL GLPLRLPPQR SSNRDPAPRG PPKPQLHFLK TSAVQYWPY

181 EKVPAFLHPF PG

Pan paniscus (pygmy chimpanzee) FGF21 (GenBank Accession No. XP_003814163, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 148); partial sequence corresponding to human FGF21 residues 1 to 116 and 195 to 201

1 MDSDETGFEH SGLWVSVLAG LLLGACQAHIP IPDSSPLLQF GGQVRQRYLY TDDAQTEAH

61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSVSF----

121 -----Q-----DPP--

181 --HHPP---C S---SYMSPS Q---PG---

TABLE 3 -continued

Macaca fascicularis (crab-eating macaque) FGF21 (GenBank Accession No. EHH59757, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 149); partial sequence corresponding to human FGF21 residues 1 to 116
1 MDSDETGFEH SGLWVPVLAG LLLGACQAHP IPDSSPLLQF GGQVRQRYLY TDDAQQTEAH

61 LEIREDTGTVG GAAHQSPESL LQLKALKPGV IQILGVKTSR FLCQKPDGAL YGSVSF

Mesocricetus auratus (golden hamster) FGF21 (GenBank Accession No. ACB30542, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 150); partial sequence corresponding to human FGF21 residues 90 to 193
1 VIQILGVKAAR FPCQQPDGS LYGSPHFDPE ACSFRELLLE DGYNVYQSEA HGLPLRLPQR

61 DAPSQPPASW GPVRFPLPVG LFQPPPHDLPG RPAPEPPDVG SSDP

Oreochromis niloticus (Nile tilapia) FGF21 (GenBank Accession No. XP_003438516, which is hereby incorporated by reference in its entirety) (SEQ ID NO: 151); partial sequence corresponding to human FGF21 residues 59 to 209
1 MYLQMNMNDGR VTGSDAQTPY SLMQLKSVKP GHVIIKGPSL SLFLCDVSEG NLRGQSHYSE

61 TSCTPREMILL ADGYTRFISS QYGFPMSLAS RHSPDRHALP FTRFLPLRNN LKTDSVSEQL

121 PNNQRLFNV DDDLLGMGLN SMGSPQFSMD K

In one embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein of the present invention comprises an amino acid residue substitution to strengthen or increase the stability of the FGF21 core domain compared to wild type FGF21. In one embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein of the present invention comprises an amino acid residue substitution to strengthen or increase the stability of the FGF21 core domain compared to that of SEQ ID NO: 100. In one particular embodiment, the N-terminal portion of FGF21 comprises a substitution at a residue corresponding to residue 104 of SEQ ID NO: 100. In one embodiment, the substitution is a glutamine to methionine substitution (i.e., Q104M).

The N-terminal portion of the chimeric protein according to the present invention may include a core domain, also

25 FGF ligands at the corresponding position (Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine & Growth Factor Rev* 16(2):107-137 (2005), which is hereby incorporated by reference in its entirety), increases the stability of FGF21 without affecting ligand-binding affinity for receptor.

30 In one embodiment of the present invention, FGF21 has an amino acid sequence corresponding to human FGF21 harboring a mutation at Q104. In one embodiment of the present invention, FGF21 has an amino acid sequence corresponding to human FGF21 harboring a Q104M mutation. In one embodiment the FGF21 having an amino acid sequence corresponding to human FGF21 and harboring a 35 Q104M mutation, has the amino acid sequence of SEQ ID NO: 152, as follows:

(SEQ ID NO: 152)

1 MDSDETGFEH SGLWVSVLAG LLLGACQAHP IPDSSPLLQF GGQVRQRYLY TDDAQQTEAH
61 LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCMRRPDGAL YGSLHFDPEA
121 CSFRELLLE GYNVYQSEAHL GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI
181 LAPQPPDVGS SDPLSMVGPQ QGRSPSYAS

50

referred to as, for example, an FGF21 core domain. In one embodiment, the core domain is the FGF β-trefoil core domain. In one embodiment, this region corresponds to H29 to L167 of human FGF21 of SEQ ID NO: 100.

In one embodiment, increasing the stability of the core domain includes an increase in thermal stability of the protein as compared to either wild type protein or a chimeric protein in which such a substitution is not made. In one embodiment, increasing the stability includes increasing the half-life of the protein in the blood circulation as compared to either wild type protein or a chimeric protein in which such a substitution is not made.

Based on the inventors' extensive knowledge of the structures of FGF ligands, including the structures of FGF19 and FGF23, Q104 of FGF21 was selected for mutagenesis. Replacing Q104 with methionine, which is found in all other

55 In one embodiment according to the present invention, the chimeric FGF21 protein comprises an N-terminal portion of FGF21 that contains at least one amino acid residue substitution to increase stability of the FGF21 core domain as compared to a sequence corresponding to SEQ ID NO: 100. In one embodiment of the present invention, the N-terminal portion of FGF21 comprises an amino acid sequence spanning residues corresponding to residues selected from the group consisting of from position 29 to 167 of SEQ ID NO: 152, from position 29 to 190 of SEQ ID NO: 152, or from position 29 to 197 of SEQ ID NO: 152. Exemplary chimeric proteins include those of SEQ ID Nos: 312-336.

60 In one particular embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein of the present invention is a modified N-terminal portion of the FGF21 protein. In one embodiment, the N-terminal portion

of the chimeric protein of the present invention comprises an amino acid sequence at least 85% identical to the amino acid sequence corresponding to residues from position 29 to 197 of SEQ ID NO: 100, from position 29 to 190 of SEQ ID NO: 100, or from position 29 to 167 of SEQ ID NO: 100. In one embodiment, the N-terminal portion of FGF21 of the chimeric protein of the present invention is derived from a modified FGF21 protein, where the N-terminal portion of the chimeric protein of the present invention comprises an amino acid sequence that has at least 80%, at least 85%, at least 90%, at least 95%, at least 97% or at least 99% amino acid sequence identity to the amino acid sequence corresponding to residues from position 29 to 197 of SEQ ID NO: 100, from position 29 to 190 of SEQ ID NO: 100, or from position 29 to 167 of SEQ ID NO: 100. In one embodiment, the N-terminal portion having such amino acid sequence similarity will maintain the activity of the corresponding naturally occurring N-terminal portion of FGF21. In one embodiment, the N-terminal portion of the chimeric protein of the present invention comprises an amino acid sequence at least 85% homologous to the amino acid sequence corresponding to residues from position 29 to 197 of SEQ ID NO: 100, from position 29 to 190 of SEQ ID NO: 100, or from position 29 to 167 of SEQ ID NO: 100. In one embodiment, the N-terminal portion of FGF21 of the chimeric protein of the present invention is derived from a modified FGF21 protein, where the N-terminal portion of the chimeric protein of the present invention comprises an amino acid sequence that has at least 80%, at least 85%, at least 90%, at least 95%, at least 97% or at least 99% amino acid sequence homology to the amino acid sequence corre-

sponding to residues from position 29 to 197 of SEQ ID NO: 100, from position 29 to 190 of SEQ ID NO: 100, or from position 29 to 167 of SEQ ID NO: 100. In one embodiment, the N-terminal portion having such amino acid sequence homology will maintain the activity of the corresponding naturally occurring N-terminal portion of FGF21.

In one embodiment of the present invention, the N-terminal portion of FGF21 of the chimeric protein of the present invention is a modified N-terminal portion of the FGF21 protein. In one embodiment, the N-terminal portion of FGF21 of the chimeric protein of the present invention comprises an amino acid sequence at least 85% identical to the amino acid sequence corresponding to residues from position 29 to 167 of SEQ ID NO: 152, from position 29 to 190 of SEQ ID NO: 152, or from position 29 to 197 of SEQ ID NO: 152. In one embodiment, the N-terminal portion of FGF21 of the chimeric protein of the present invention comprises an amino acid sequence at least 85% homologous to the amino acid sequence corresponding to residues from position 29 to 167 of SEQ ID NO: 152, from position 29 to 190 of SEQ ID NO: 152, or from position 29 to 197 of SEQ ID NO: 152.

It will be understood that the portion of FGF21 of the chimeric protein of the present invention may be derived from a nucleotide sequence that encodes a vertebrate or a non-vertebrate FGF21 protein. In one embodiment, the portion of FGF21 of the chimeric protein of the present invention may be derived a nucleotide sequence that encodes a mammalian FGF21 protein. Nucleotide sequences encoding a vertebrate FGF21 protein according to the present invention may include, but are not limited to, those shown in Table 4.

TABLE 4

Human FGF21 gene coding sequence (SEQ ID NO: 153) (GenBank Accession No. NM_019113, which is hereby incorporated by reference in its entirety)	
151	ATGGAACCGG ACGAGACCGG GTTCGAGCAC TCAGGACTGT GGGTTCTGT GCTGGCTGGT
211	CTTCTGCTGG GAGCCTGCCA GGCACACCCC ATCCCTGACT CCAGTCCTCT CCTGCAATT
271	GGGGGCCAAC TCGGCAGCG CTACCTCTAC ACAGATGATG CCCAGCAGAC AGAAGCCAC
331	CTGGAGATCA GGGAGGATGG GACGGTGGGG GGCGCTGCTG ACCAGAGCCC CGAAAGTCTC
391	CTGCAGCTGA AAGCCTTGAA GCCGGGAGTT ATTCAAATCT TGGGAGTC GACATCCAGG
451	TTCCTGTGCC AGCGGCCAGA TGGGGCCCTG TATGGATCGC TCCACTTGA CCCTGAGGCC
511	TGCAGCTTCC GGGAGCTGCT TCTTGAGGAC GGATAACAATG TTTACCAGTC CGAACGCCAC
571	GGCCTCCCGC TGCACCTGCC AGGAAACAAG TCCCCACACC GGGACCCCTGC ACCCCGAGGA
631	CCAGCTCGCT TCCTGCCACT ACCAGGCCTG CCCCCCGCAC TCCCGGAGCC ACCCGGAATC
691	CTGGCCCCCGA AGCCCCCGA TGTGGCTCC TCGGACCTCT TGAGCATGGT GGGACCTTC
751	CAGGGCCGAA GCCCCAGCTA CGCTTCCCTGA
<i>Pongo abelii</i> (Sumatran orangutan) FGF21 gene coding sequence (SEQ ID NO: 154) (GenBank Accession No. XM_002829519, which is hereby incorporated by reference in its entirety)	
165	ATGGAC TCGGACGAGA CGGGGTCGA GCACTCAGGA CTGTGGTTC CTGTGCTGGC
221	TGGCTTCTG CTGGGAGCCT GCCAGGCACA CCCCATCCCT GACTCCAGTC CTCTCCTGCA
281	ATTCGGGGC CAAGTCCGGC AGCGGTACCT CTACACAGAT GATGCCAGC AGACAGAAGC
341	CCACCTGGAG ATCAGGGAGG ATGGGACGGT GGGGGCGCT GCTGACCAGA GCCCCGAAAG
401	TCTCCTGCAG CTGAAAGCCT TGAAGCCGGG AGTTATTCAA ATCTTGGGAG TCAAGACATC
461	CAGGTTCCCTG TGCCAGAGGC CAGATGGGC CCTGTATGGA TCGCTCCACT TTGACCCCTGA

TABLE 4-continued

521 GGCCTGCAGC TTCCGGGAGC TGCTTCTTGA GGACGGATAC AATGTTTATC AGTCCGAGGC
 581 CCATGCCCTC CCGCTGCACC TGCGGGAAA CAAGTCCCCA CACCGGGACC CTGCACCCCG
 641 AGGACCAGCT CGCTTCCTGC CACTACCAGG CCTGCCCCCC GCACCCCCAG AGCCGCCCG
 701 AATCCTGGCC CCCAGCCCC CCGATGTGGG CTCCCTGGAC CCTCTGAGCA TGGTGGGACC
 761 TTCCCAGGGC CGAAGCCCCA GCTATGCTTC CTGA

Pan troglodytes (chimpanzee) FGF21 gene coding sequence
 (SEQ ID NO: 155) (GenBank Accession No. XM_524333, which
 is hereby incorporated by reference in its entirety)

573 ATGGACTC GGACGAGACC GGGTCGAGC ACTCAGGACT GTGGGTTCT GTGCTGGCTG
 631 GTCTTCTGCT AGGAGCCTGC CAGGCACACC CCATCCCTGA CTCCAGTCCT CTCCCTGCAAT
 691 TCGGGGGCCA AGTCCGGCAG CGGTACCTCT ACACAGATGA TGCCCCAGCAG ACAGAACGCC
 751 ACCTGGAGAT CAGGGAGGAT GGGACGGTGG GGGCGCTGC TGACCAGAGC CCCGAAAGTC
 811 TCCTGCAGCT GAAAGCCTTG AAGCCGGAG TTATTCAAAT CTTGGGAGTC AAGACATCCA
 871 GGTTCTGTG CCAGAGGCCA GATGGGGCCC TGTATGGATC GCTCCACTTT GACCCCTGAGG
 931 CCTGCAGCTT CGGGAGCTG CTTCTTGAGG ACGGATAACA TGTTTACCAAG TCCGAGGCC
 991 ACGGCCTCCC GCTGCACCTG CCGGGAAACA AGTCCCCACA CCGGGACCT GCACCCCGAG
 1051 GACCAGCTCG CTTCTGCCA CTACCAGGCC TGCCCCCGC ACCCCCGGAG CCACCCGGAA
 1111 TCCTGGCCCC CCAGCCCCCC GATGTGGGCT CCTCAGACCC TCTGAGCATG GTGGGACCTT
 1171 CCCAGGGCCG AAGCCCCAGC TACACTTCCT GA

Canis lupus familiaris (dog) FGF21 gene coding sequence
 (SEQ ID NO: 156) (GenBank Accession No. XM_541510, which
 is hereby incorporated by reference in its entirety)

1 ATGGGCTGGG CCGAGGCCGG GTTCGAGCAC CTGGGACTGT GGGTCCCTGT GCTGGCTGTG
 61 CTTTGCTGG AAGCCTGCCG GGCACATCCG ATCCCTGACT CCAGCCCCCT CCTACAATTT
 121 GGAGGTCAAG TTCGACAGCG GTACCTCTAC ACCGACGATG CCCAGGAGAC AGAGGCCAC
 181 CTAGAGATCA GGGCCGATGG CACAGTGGTG GGGCCTGCCG GCCAGAGCCC TGAAAGTCTC
 241 CTGGAGCTGA AAGCCCTAAA GCCAGGGGTC ATTCAAATCT TGGGAGTCAA AACATCCAGG
 301 TTCCTGTGCC AGGGCCAGA TGGGACACTA TATGGCTCGC TCCATTTCGA CCCTGTGGCC
 361 TGCAGTTCC GAGAACTGCT TCTTGAGGAT GGGTACAACA TCTACCACTC CGAGACCCCTT
 421 GGTCTCCCGC TTCGCTGCG CCCCCACAAC TCCGCATACC GGGACTTGGC ACCCCCGGG
 481 CCTGCCCGCT TCCTGCCACT GCCAGGCCTG CTTCCAGCAC CCCCAGAGCC TCCAGGGATC
 541 CTGGCCCCGG AGCCTCTGA CGTGGGCTCC TCGGACCCCTC TGAGCATGGT GGGGCCTTC
 601 CAGGGCCCGA GTCCCAGCTA TGCTTCCTAA

Bos taurus (bovine) FGF21 gene coding sequence (SEQ ID
 NO: 157) (GenBank Accession No. XP_001789587, which is
 hereby incorporated by reference in its entirety)

1 ATGGGCTGGG ACGAGGCCAA GTTCAGCAC TTGGGACTGT GGGTCCCTGT GCTGGCTGT
 61 CTCCCTGTAG GAAACCTGCCG GGCACATCCC ATTCCAGACT CCAGCCCCCT CCTCCAGTTT
 121 GGGGGCCAAG TCCGCCAGCG GTACCTCTAC ACGGATGATG CCCAGGAGAC AGAGGCCAC
 181 CTGGAGATCA GGGCCGATGG CACAGTGGTG GGGCAGGCCG GCCAGAGCCC CGAAAGTCTC
 241 TTGGAGCTGA AAGCCCTGAA GCCAGGCCTG ATTCAAGATCT TGGGAGTTAA AACATCCAGG
 301 TTTCTCTGCC AGGGCCAGA TGGGAGCTG TACGGATCGC TGCACATTGAA CCCCCAAAGCC
 361 TGCAGCTTTC GGGAGCTGCT TCTTGAAGAT GGATACAACG TCTACCAAGTC GGAGACCCCTG
 421 GGCCTTCCAC TCCGCCCTGCC CCCCCAGCGC TCGTCCAACC GGGACCCGGC CCCGCCGGGA

TABLE 4-continued

481 CCTGCTCGCT TCCTTCACT GCCGGGCCTG CCCGGCGCGC CCCCAGATCC TCCAGGGATC
 541 TTGGCCCCCG AGCCTCCGA CGTGGGCTCC TCGGATCCCC TGAGTATGGT GGGACCCCTCG
 601 TATGGCCGAA GCCCCAGCTA CACTTCTTGA

Equus caballus (horse) FGF21 gene coding sequence (SEQ ID NO: 158) (GenBank Accession No. XM_001489152, which is hereby incorporated by reference in its entirety)

1 ATGGACTGGG ACAAGACGGG GTTCAGTAC CAGGGACTGT GGGTCCCTGT GCTGGCTGTC
 61 CTTCTGCTGG GAGCCTGCCA GTCACACCCCC ATCCCTGACT CCAGTCCCC CCTCCAATT
 121 GGGGGCCAAG TCAGGCAGCG CCACCTCTAC ACAGATGATG CCCAGGAGAC AGAGGCGCAC
 181 CTGGAGATCA GGGCTGACGG CACTGTGGCA GGGGCTGTCC ACCGGAGCCC AGAAAGTCTC
 241 TTGGAGCTGA AAGCCCTGAA GCCAGGGTA ATTCAAATCT TGAGTCAA GACATCCAGG
 301 TTCCTGTGCC AGGGGCCAGA CGGGACGCTG TACGGATCGC TCCACTCGA CCCCCTGGCC
 361 TGCAGCTTCC GGGAGCTGCT TCTCGAAGAC GGCTACAACG TTTACCAGTC TGAGACCCCT
 421 GGCCTCCAC TCCGCTGCC CCACACAGC TCCCCATACC AGGATCCGGC CCCTCGGGCA
 481 CCCGCCCCGCT TCCTGCCGCT GCCAGGCTTT CCCCCAGCAC CCCCAGGAGCC TCCAGGGATC
 541 CGGGCCCCCG AGCCCCCGGA CGTGGGCTCC TCGGACCCCC TGAGCATGGT GGGGCTTCA
 601 CGCAGCCGAA GCCCCAGCTA CACTTCTTGA

Ailuropoda melanoleuca (giant panda) FGF21 gene coding sequence (SEQ ID NO: 159) (GenBank Accession No. XM_002917864, which is hereby incorporated by reference in its entirety)

1 ATGGGCTGGG ACGAGGCCAG GTCCGAGCAG CTGGGGCTGT GGGTCCCTGT GCTGGCTGTC
 61 CTTTGCTGG AAGCTTGCCTA GGCACACCCCT ATCCCTGACT CCAGTCCCC CCTCCAATT
 121 GGAGGCCAAG TTGACAGCG GTACCTCTAC ACGGACGATG CCCAGGAGAC AGAGGCCAC
 181 CTAGCGATCA GGGCTGATGG CACAGTGGTG GGGGCTGCCA GCCGGAGCCC AGAAAGTCTC
 241 TTGGAGCTGA AAGCCCTGAA ACCGGGGTC ATTCAAATCC TGAGTCAA AACATCTAGG
 301 TTCCTGTGCC AGGGGCCAGA TGGGACACTG TACGGATCGG TCCGCTCGA CCCCCTAGCC
 361 TGCAGCTTCC GGGAACTGCT CCTGGAGGAT GGGTACAACA TCTACCACTC TGAGACCCCTC
 421 GGCCTCCAC TCCGCTGCC CGCCACAAAC TCTCCATACC GGAACTCGGC GCCCCGGGG
 481 CCTGCCGCT TCCTGCCCT GCCAGGCCTG CTTCCGGTCC CCCCCGACCC CCCAGGGATC
 541 CTGGGCCCCG AGCCTCCGA CGTGGGCTCC TCGGACCCCC TGAGCATGGT GGGGCTTCA
 601 CAGGGCCGAA GTCCCAGCTA CGCTTCTTGA

Oryctolagus cuniculus (rabbit) FGF21 gene coding sequence (SEQ ID NO: 160) (GenBank Accession No. XM_002723699, which is hereby incorporated by reference in its entirety)

1 ATGGACTGGG GCAAGGCCAA GTGCCGGCCCC CGGGGGCTGT GGGTCCCTGC GCTCGCTGCC
 61 CTGCTGCTGG GGGCCTGCCA GGCACACCCCC ATCCCGACT CCAGTCCCC CCTCCAGTT
 121 GGGGACCAAG TGCAGCAGCA GCACCTGTAC ACGGACGATG CGCAGGAAAC AGAAGGCCAC
 181 CTGGAGATCA GGGCGGATGG CACGGTGGTG GGGGCTGCCA GGAGGAGCCC AGAAAGTCTC
 241 TTGCAGATGA AAGCCTTACA ACCGGGGATC ATTCAAGATCT TGAGGGTCCA GACGTCCAGG
 301 TTCCTCTGCC AGAGGCCGA TGGCACGCTC TACGGCTCGC TCCACTCGA CCGCGAGGCC
 361 TGCAGCTTCC GGGAGCTGCT GCGTGGAGAT GGGTACAACG TTTACCTCTC GGAGGCCCTG
 421 GGCCTGCCCT TGCGCTGTC CCCCGCGAGC TCCCCACGCA GGGCGCCGGC CCCCCGGGG
 481 CCAGCCCCGCT TCCTGCCGCT GCCCGGCCTG CGCCAGAACCC TTCCGGAACC GCCAGGCCCTC

TABLE 4-continued

541 CTGGCCGCCG CGCCCCCGA TGTCGACTCC CCGGACCCCC TGAGCATGGT GCAGCCTGCG

601 CTGGACCAGA GCCCCAGCTA CACCTCCTGA

Gorilla gorilla (gorilla) FGF21 gene coding sequence
(SEQ ID NO: 161) (Ensembl Accession No. ENSGGOT00000001253,
which is hereby incorporated by reference in its entirety)

151 ATGGACTCGG ACGAGACCGG GTTCGAGCAC TCAGGACTGT GGTTTCTGT GCTGGCTGGT

211 CTTCTGCTGG GAGCCTGCCA GGCACACCCC ATCCCTGACT CCAGTCCTCT CCTGCAATT

271 GGGGGCCAAG TCCGGCAGCG GTACCTCTAC ACAGATGATG CCCAGCAGAC AGAACGCCAC

331 CTGGAGATCA GGGAGGATGG GACGGTGGGG GGTGCTGCTG ACCAGAGCCC TGAAAGTCTC

391 CTGCAGCTGA AAGCCTGAA GCCGGGAGTT ATTCAAATCT TGGGAGTCAA GACATCCAGG

451 TTCCTGTGCC AGAGGCCAGA TGGGGCCCTG TATGGATCGC TCCACTTTGA CCCTGAGGCC

511 TGCAGCTTCC GGGAGCTGCT TCTTGAGGAC GGATACAATG TTTACCAGTC CGAGGCCAC

571 GGCCTCCCGC TGCACCTGCC GGGAAACAAG TCCCCACACC GGGACCCCTGC ACCCCGAGGA

631 CCAGCTCGCT TCCTGCCACT ACCAGGCCCTG CCCCCCGCAC CCCCCGGAGCC ACCCGGAATC

691 CTGGCCCCCC AGCCCCCGA TGTGGGCTCC TCGGACCCCTC TGAGCATGGT GGGACCTTCC

751 CAGGGCCGAA GCCCCAGCTA CGCTTCCCTGA

Nomascus leucogenys (Northern white-cheeked gibbon) FGF21
gene coding sequence (SEQ ID NO: 162) (Ensembl Accession
No. ENSNLET00000005931, which is hereby incorporated by
reference in its entirety)

587 ATGG ACTCGGACGA GACCGGGTTC GAGCACTCG GACTGTGGGT TCCTGTGCTG

647 GCTGGCTTC TGCTGGGAGC CTGCCAGGCA CACCCATCC CTGACTCCAG TCCTCTCCTG

707 CAATTGGGG GCCAAGTCCG GCAGCGGTAC CTCTACACAG ATGATGCCA GCAGACAGAA

767 GCCCACCTGG AGATCAGGGA GGATGGGACG GTGGGGGGCG CTGCTGACCA GAGCCTGAA

831 AGTCTCTGC AGCTGAAAGC CTTGAAGCCG GGAGTTATTC AAATCTTGGG AGTCAAGACA

891 TCCAGGTTCC TATGCCAGAG GCCAGATGGG GCCCTGTATG GATCGCTCCA CTTTGACCC

951 GAGGCCTGCA GCTTCCGGGA GCTGCTTCTT GAGGACGGAT ACAATGTTA CCAGTCCGAG

1011 GCCCCATGGCC TCCCGCTGCA CCTGCCGGGG AACAACTCCC CACACCGGGA CCCTGCACCC

1071 CGAGGACCAAG CTCGCTTCCT GCCACTACCA GGCCCTGCCCT CTGCACCCCG AGAGCCGCC

1131 GGAATCCTGG CCCCCCAGCC CCCCGATGTG GGCTCCTCGG ACCCTCTGAG CATGGTGGGA

1191 CCTTCCCCAGG GCCGAAGCCC CAGCTACGCT TCCTGA

Procavia capensis (hyrax) FGF21 gene coding sequence (SEQ
ID NO: 163) (Ensembl Accession No. ENSPCAT00000001288,
which is hereby incorporated by reference in its entirety)

1 ATGGACTGGG CCAAGTTGG GATCGAGCAC CGGGACTGT GGTCGGCGGT GATGGCAGTA

61 CTTCTGCTGG GAGCCTGCCA AGGATAACCT ATTCCCTGACT CCAGCCCCCT TCTCCAATT

121 GGAGGCCAGG TCCGGCAACG TTACCTCTAC ACAGATGACG CGCAGGAGAC CGAGGCCAC

181 CTGGAGATCC GAGCAGACGG CACGGTGGTG GGGCTGCC ACCGGAGCCC CGAGAGTCTC

241 TTGGAGCTGA AAGCTTGAA GCCCGGCATA ATTCAAGATCT TGGGAGTCAA GACATCCAGA

301 TTCCTCTGCC AGGGTCTGCA TGGGGTGTG TATGGATCGC TCCGTTTGA CCCAGTGGCC

361 TGCAAGCTCC GGGAGCTGCT TCTTGAAAGAT GGATACAATG TTTACCAGTC TGAGGCCAC

421 GGCCTCCCGC TTGCGCTTAC ATCCCACAAT TCCCCACAGA GGGACCTGGC GTCCCGGGTG

481 CCAGCCCCGT TCCTGCCACT GCCAGGCCGG CTCACGGTGC TCCAGAACCC TTCGGGGGTC

541 CTGGGCCCTG AGCCCCCGA TGTGGACTCC TCAGACCCCC TGAGCATGGT GGGACCTTCC

601 CAGGGCCGAA GCCCCAGTTA CGCCTCCTGA

TABLE 4-continued

Cavia porcellus (guinea pig) FGF21 gene coding sequence
(SEQ ID NO: 164) (Ensembl Accession No. ENSCPOT00000000273,
which is hereby incorporated by reference in its entirety)

```

1 ATGGACTGGG CCCGGACTGA GTGTGAGCGC CCAAGGCTGT GGGCTCCAT GCTGGCCATC

61 CTTCTGGTGG GAGCCTGCCA GGCACACCCT ATCCCTGACT CCAGCCCCCT CCTCCAGTTT

121 GGGGGCCAGG TCCGGCAGCG GTACCTCTAC ACAGATGATG CTCAGGACAC TGAAGTGCAC

181 CTGGAGATCA GGGCGATGG CTCAGTACGG GGCATTGCC ACAGGGGCC TGAAAGTCTC

241 CTGGAGCTGA AAGCCTGAA GCCAGGAGTC ATTCAAGATCT TGGGAATCAG GACTTCCAGG

301 TTCCTGTGCC AGAGGCCGA TGGGAGTCTG TATGGATCAC TCCACTTGAA TCCTGAGGCC

361 TGCAGCTTCC GGGAGCTGCT GCTTGCTGAT GGCTACAATG TCTACAAGTC TGAAGCCCAC

421 GGCCTCCCTC TGCACCTGCT GCGCGGTGAC TCTCTATCGC AGGAACCAGC ACCCCCCAGGA

481 CCAGCCCGAT TTCTGCCACT ACCAGGCCTG CCCGCAACAC CCCCCGGAGCC ACCCAGGATG

541 CTGCCCCCAG GGCCCCCAGA TGTGGGCTCC TCGGACCTT TGAGCATGGT GGGGCCTTTA

601 TGGGACCGAA GCCCCAGCTA TACTTCCCTGA

```

Tupaia belangeri (tree shrew) FGF21 gene coding sequence
(SEQ ID NO: 165) (Ensembl Accession No. ENSTBET00000016056,
which is hereby incorporated by reference in its entirety)

```

1 ATGGGCTGGG ACAAGGCCCG GTTCGAGCAC CTGGGAGCGT GGGCTCCTGT GCTGGCTGTC

61 CTCCCTCTGG GAGCCTGCCA GGCATAACCC ATCCCTGACT CCAGCCCCCT CCTACAATTG

121 GGGGGCCAGG TCCGGCAGCG GTACCTCTAC ACGGACGACA CGCAGGACAC AGAACGCCAC

181 CTTGAGATCA GGGCGACGG CACCGTGGTG GGGGCCGCC ACCAAAGCCC GGAAAGTCTC

241 CTGGAGCTGA AAGCCTGAA GCCGGGGGTC ATTCAAATCC TGGGAGTCAA GACCTCCAGG

301 TTCCTGTGCC AGAGGCCAGA CGGGGCCCTG TACGGGTGCG TTCACTCGA CCCCCGAGGCC

361 TGCAGCTTCC GGGAGCTGCT TCTCGAGGAT GGATACAACA TTTACCAGTC TGAGGCTCGT

421 GGCCTCCCCC TGCACCTGCC GCCCCACGAC TCCCCACATC GGGACCGGAC CCCTCGGGGA

481 CCAGCTCGTT TCCTGCCGCT GCCTGCCCTG CCCCTGGTTC CTCCAGAGCT GCCAGGGGTC

541 CTGGCCCTTG AGCCCCCAGA CGTGGGCTCC TCAGACCGC TGA

```

Sorex araneus (shrew) FGF21 gene coding sequence (SEQ ID
NO: 166) (Ensembl Accession No. ENSSART0000003074,
which is hereby incorporated by reference in its entirety)

```

1 ATGGTCTGGG ACAAGGCCAG GGGCAGCAG TTGGGACTGT GGGCCCCCAT GCTGCTGGC

61 TTGCTGCTGG GTGCCTGCCA GGCACACCCCC CTCCCTGACT CCAGCCCCCT CCTCCAATTG

121 GGGGGCCAAAG TCCGACTGAG GTTCTGTAC ACCGACGATG CCCAGAGGAC AGGGGCGCAC

181 CTGGAGATCA GGGCGACGG CACAGTGCAG GGTGCGGCC ACAGGACCCC AGAATGTCTC

241 CTGGAGCTGA AAGCCTGAA GCCAGGCAGTA ATTCAAATCC TTGGGGTCAG CACATCCAGA

301 TTCCTGTGCC AGCGGCCGA TGGGGTCCTG TATGGATCGC TTGCGTTGAA CCCAGAGGCC

361 TGCAGTTCC GGGAACTTCT TCTCCAGGAT GGATATAACG TTTACCAGTC TGAGGCCCTG

421 GGTCTCCCGC TCTACCTACA CCCGCCAGT GCCCCAGTGT CCCAGGAACC AGCCTCACGG

481 GGCGCCGTCC GCTTCCTGCC ACTGCCAGGA CTGCCACCTG CCTCCCTGGA GCCCCCCAGG

541 CCCCCCGCCC CGGTGCCTCC AGACGTGGGT TCCTCAGACC CCCTGA

```

Ictidomys tridecemlineatus (squirrel) FGF21 gene coding
sequence (SEQ ID NO: 167)

```

1 ATGTACCCCA TCCCTGACTC AAGCCCCCTC CTCCAATTG GGGGCAAGT CCCGGCAGCG

61 TACCTGTACA CAGATGATGC CCAGGAGACT GAGGCCAC TGGAGATCAG GGCTGATGGC

121 ACCGTGGTGG GGGCTGCCCA TCAAAGCCCG GAAAGTCTCT TGGAACTGAA AGCCTGAAG

```

TABLE 4-continued

181 CCTGGGGTCA TTCAAATCTT GGGGGTCAAACATCCAGGT TCCTGTGCCA GAGGCCAGAT
 241 GGAGTGCTGT ATGGATCGCT CCACTTGAC CCTGAGGCCT GCAGCTCCG GGAGCAGCTT
 301 CTGGAGGACG GGTACAACGT TTACCAGTCA GAATCCCACG GCCTCCCCGT GCGCCTGCC
 361 CCTAACTCAC CATAACGGGA CCCAGCGCCG CCAGGACAG CCCGCTTCCT TCCACTGCCA
 421 GGCCCTGCCCT CAGCAGCCCT GGAGCCGCCA GGATCCTGG GCCTGAGCC CCCTGATGTG
 481 GGCTCCTCCG ACCCACTCAG CATGGTGGGG CCTTGAGG GCCGAAGCCC CAGTTACGCT
 541 TCCTGA

Loxodonta africana (elephant) FGF21 gene coding sequence
 (SEQ ID NO: 168) (Ensembl Accession No. ENSLAFT00000022429,
 which is hereby incorporated by reference in its entirety)
 1 ATGGACTGGG CCAAGTTGG GTTGGAGCAC CCAGGACTGT GGGTCCCTGT GATGGCTGTC

61 CTTCTGCTGG GAGCCTGCCA GGGACACCCCC ATCCCTGACT CCAGCCCCCT CCTCCAATT
 121 GGGGGCCAGG TCCGGCAACG TTACCTCTAC ACAGATGATC AGGAGACCGA GGCCCACCTG
 181 GAGATCAGAG CAGATGGCAC AGTGGCGGGA GCCGCTCACC GGAGCTCTGA GAGTCTCTTG
 241 GAGCTGAAAG CTTTGAAGCC TGGAATAATT CAGATCTTGG GGGTCAAGAC ATCCGGTTC
 301 CTGTGCCAGG GGCCTGATGG GGTGCTGTAC GGATCGCTCC ATTCGACCC AGCCGCTGC
 361 AGCTTCCGGG AGCTGTTCT TGAAGATGGA TACAATGTTT ACTGGTCCGA GGCCCATGGA
 421 CTCCCAATCC GCCTGCCCTC CCACAACCTCC CCATATAGGG ACCCAGCATC CGGGGTACCA
 481 GCCCGCTTCC TGCCACTGCC AGGCCTGCTC CCAATGCTCC AAGAACCTCC AGGGGTCTG
 541 GCCCCTGAGC CCCCTGATGT GGACTCCTCA GACCCCTGA GCATGGTGGG GCCTTCACAG
 601 GGCCGAAGCC CCAGCTATGC CTCCCTGA

Sus scrofa (pig) FGF21 gene coding sequence (SEQ ID NO: 169) (GenBank Accession No. NM_001163410, which is hereby incorporated by reference in its entirety)
 131 ATGGGCTGGG CCGAGGCCAA GTTCGAGGCC TTGGGACTGT GGGTCCCTGT GCTGGCTGTC

191 CTGCTGGGAG CCTGCCAGGC ACGTCCCATT CCTGACTCCA GCCCCCTCCT CCAATTGGG
 251 GGCCAAGTGC GCCAACGATA CCTCTACACG GATGATGCC AGGAAACTGA AGCCCACCTG
 311 GAGATCAGAG CTGATGGCAC CGTGGCAGGG GTAGCCCGCC AGAGCCCTGA AAGTCTCTTG
 371 GAGCTGAAAG CCCTGAAGCC AGGGGTCAATT CAAATTTGG GAGTCCAGAC ATCCGGTTC
 431 CTGTGCCAGG GGCCAGACGG GAGACTGTAC GGATCGCTCC ACTTCGACCC TGAGGCCTGC
 491 AGCTTCCGGG AGCTGTTCT TGAGGATGGC TACAACGTTT ACCAGTCTGA GGCCCTTGGC
 551 CTCCCACTCC GGCTGCCCTC GCACCGCTCC TCCAACCGGG ACCTGGCCCG CGGGGACCT
 611 GCTCGCTTCC TGCCACTGCC AGGCCTGCC CCGGCACCCCC CGGAGCCGCC AGGGATCTTG
 671 GCCCCTGAAC CTCCCGACGT GGGCTCCTCG GACCCCTGA GCATGGTGGG GCCTTCACAC
 731 GGCCGGAGCC CCAGCTACAC TTCTTG

Felis catus (cat) FGF21 gene coding sequence (SEQ ID NO: 170) (Ensembl Accession No. ENSFCAT00000007367, which is hereby incorporated by reference in its entirety)
 1 ATGGGCTGGG ACGAGGCCGG GTCCAGCGC CTGGGACTGT GGGTCTGCT GGGGGTCTT

61 TTGCCGGAAG CCTGCCAGGC ACACCCATC CCTGACTCCA GCCCCCTCCT CCAATTGGG
 121 GGCCAAGTTC GACAGCGGTT CCTCTACACG GACGACGCC AGGAGACAGA GGTCCACCTC
 181 GAGATCAAGG CTGATGGCAC AGTGGTGGGG ACCGCTGCC GGAGCCCTGA GAGTCTCTTG
 241 GAGCTAAAG CCCTGAAGCC GGGGTAATT CAAATCTTGG GGGTCAAAAC GTCCAGGTTC
 301 CTGTGCCAGG GCCCAGATGG GACACTGTAT GGATCGCTCC GCTTGACCC CGCAGCCTGC
 361 AGCTTCCGGG AACTGCTCCT GGAGGACGGA TACAACATCT ACCACTCGGA GACCCCTCGGG

TABLE 4-continued

421 CTCCCCACTCC GCCTGCCCCC CCACAACTCC CCATAACCGGG ACTTGGCCCC CGGGCACCT
 481 GCCCGCTTCC TGCCGCTGCC AGGCCTGTT CCGGCACCCC CGGAGCCTCC AGGGATCCTG
 541 GCCCCCGAGC CCCCGGACGT GGGCTCCTCG GACCCTCTGA GCATGGTGGG GCCTTCCCAG
 601 GGCGAAGTC CCAGCTACGC TTCCTGA

Otolemur garnettii (bushbaby) FGF21 gene coding sequence
 (SEQ ID NO: 171) (Ensembl Accession No. ENSOGAT00000003585,
 which is hereby incorporated by reference in its entirety)
 1 GACAAGGCCA GGACTGGGTT CAAGCACCA GGACCATGGT TTCCCCTGCT GGCTGTACTT

61 TTGTTGGGAG CCTGCCAGGC ACACCCCTATC CCTGACTCCA GCCCCCTACT CCAGTTGGT
 121 GGCCAAGTCC GGCAGCGGTA CCTCTACACA GATGATGCC AGGAGACAGA AGCCCACCTG
 181 GAGATCAGGG AAGATGGCAC AGTGGTGGGG GCTGCACAAAC AGAGCCCTGA AAGTCTCTTG
 241 GAGCTGAAAG CTTAAAGCC AGGGGTCAATT CAAATCTTGG GAGTCAAGAC ATCCAGGTTC
 301 CTGTGCCAGA GGCCAGATGG GGGCTATAT GGATCGCTCT ACTTTGACCC CAAGGCCTGC
 361 AGTTTCCGGG AGCTGCTTCT TGAGGATGGA TACAACGTTT ACTGGCTGTA GACCTATGGC
 421 CTCCCCACTGC ACCTGCCTCC TGCCAATTCC CCATACTGGG GCCCATCCCT TCGGAGCCCA
 481 GCCCGCTTCC TGCCACTGCC AGGCCCTCCT GCAGCATCCC CAGAGCTGCC GGGGATCTTG
 541 GCCCTGGAAC CCCCGATGT GGGCTCCTCG GACCCTCTGA GCATGGTGGG GCCTTCGCAG
 601 GGCGAAGCC CCAGCTATGC TTCCTGA

Rattus norvegicus (Norway rat) FGF21 gene coding sequence
 (SEQ ID NO: 172) (GenBank Accession No. NM_130752, which
 is hereby incorporated by reference in its entirety)
 1 ATGGACTGGA TGAAATCTAG AGTTGGGGCC CGGGGACTGT GGGTCTGTCT CCTGCTGCCT

61 GTCTTCTGC TGGGGGTGTG CGAGGCATAC CCCATCTCTG ACTCCAGCCC CCTCCTCCAG
 121 TTTGGGGGTC AAGTCGACA GAGGTATCTC TACACAGATG ACGACCAGGA CACCGAAGCC
 181 CACCTGGAGA TCAGGGAGGA CGAACAGTG GTGGGCACAG CACACCGCAG TCCAGAAAGT
 241 CTCCCTGGAGC TCAAAGCCTT GAACCCAGGG GTCATTCAAAC TCTGGGTGT CAAAGCCTGT
 301 AGGTTCTTT GCCAACAACC AGATGGAAC CTCTATGGAT CGCCTCACTT TGATCCTGAG
 361 GCCTGCAGTT TCAGAGAGCT GCTGCTTAAG GACGGATACA ATGTGTACCA GTCTGAGGCC
 421 CATGGCCTGC CCCTGCCTCT GCCCCAGAAG GACTCCCAGG ATCCAGCAAC CCGGGGACCT
 481 GTGCGCTTCC TGCCCATGCC AGGCCCTGCC CACGAGCCCC AAGAGCAACC AGGAGTCCTT
 541 CCCCGAGAGC CCCCGATGT GGGTCCCTCC GACCCCTCTGA GCATGGTAGA GCCTTGCAA
 601 GGCGAAGCC CCAGCTATGC ATCTTGA

Mus musculus (house mouse) FGF21 gene coding sequence
 (SEQ ID NO: 173) (GenBank Accession No. NM_020013, which
 is hereby incorporated by reference in its entirety)
 185 ATGGAA TGGATGAGAT CTAGAGTTGG GACCCCTGGGA CTGTGGGTCC GACTGCTGCT

241 GGCTGCTTCC CTGCTGGGGG TCTACCAAGC ATACCCCATC CCTGACTCCA GCCCCCTCCT
 301 CCAGTTGGG GGTCAAGTCC GGCAGAGGTA CCTCTACACA GATGACGACC AAGACACTGA
 361 AGCCCCACTG GAGATCAGGG AGGATGGAAC AGTGGTAGGC GCAGCACACC GCAGTCCAGA
 421 AAGTCTCCTG GAGCTCAAAG CCTTGAAGCC AGGGTCATT CAAATCCTGG GTGTCAAAGC
 481 CTCTAGGTTT CTTTGCCAAAC AGCCAGATGG AGCTCTCTAT GGATCGCTC ACTTTGATCC
 541 TGAGGCTGC AGCTTCAGAG AACTGCTGCT GGAGGACGGT TACAATGTGT ACCAGTCTGA
 601 AGCCCATGGC CTGCCCTGC GTCTGCCTCA GAAGGACTCC CCAAACCCAGG ATGCAACATC
 661 CTGGGGACCT GTGCGCTTCC TGCCCATGCC AGGCCCTGCTC CACGAGCCCC AAGACCAAGC

TABLE 4-continued

721 AGGATTCTCG CCCCCAGAGC CCCCAGATGT GGGCTCCTCT GACCCCTGA GCATGGTAGA

781 GCCTTTACAG GGCGAAGCC CCAGCTATGC GTCTGAG

Vicugna pacos (alpaca) FGF21 gene coding sequence (SEQ ID NO: 174) (Ensembl accession no. ENSVPAT0000005993, which is hereby incorporated by reference in its entirety) (1-209, excluding 79-168 and 172-182)

1 ATGGACTGGG ACGAGGCCAA GTTCGAGCAT CGGGGACTGT GGGTCCCAGT GCTCACTGTC

61 CCTCTGCTGG GAGCCTGCCA GGCACGCCCT ATTCTGACT CCAGCCCCCT CCTCCAATTG

121 GGGGGCCAAG TCCGGCAGCG GTACCTCTAC AC GGATGACG CCCAGGAGAC AGAAGCCCAC

181 CTGGAGATCA GGGCTGATGG CACAGTGGTG GGGTGGCCC GCCAG---CC CGAA-----

241 -----

301 -----

361 -----

421 -----

481 ----- GGAATT CCT-----

541 -----CCCG AGCCTCCTGA CGTGGGCTCC TCAGACCCCC TGAGCATGGT GGGGCCTTCA

601 TACAGCAGAA GCCCCAGCTA CACTTCCTGA

Anolis carolinensis (anole lizard) FGF21 gene coding sequence (SEQ ID NO: 175) (Ensembl accession no. ENSACAT00000017230, which is hereby incorporated by reference in its entirety)

1 TGTAAGACAGGGAGG GAAGGGGGGA GAGAGGATGT GGGTAGACCT AGTTTCTGG

61 GCTGCCTTGC TCCGCACAGC TCCTGCTCTT CCCTGCGGA ATTCCAACCC CATCTACCAA

121 TTGATGGGAGGTCCCGTCTGGGCTTCAGCACCTC TACACAGCAG ATGAACAGAC GCACCTCCAC

181 TTGGAGATCT TGCCAGACGG TACCGTGGGT GGATCCAGGT TTCAGAACATCC CTTCAGTTTG

241 ATGGAGATCA AAGCTGTGAA GCCAGGAGTC ATTCCGATGC AGGCAAGAA GACCTCTAGA

301 TTCTCTGTATGAAACCCAA TGGACGACTG TATGGCTCGC TGTTCTACTC TGAGGAGGCA

361 TGCAACTTCC ATGAGAAGGT TCTCAGCGAT GGCTACAACC TCTACTATTC TGAAAATAC

421 AACATACCTG TCAGCCTCAG CTGGCAGGG AACCTGGGTC AGAGCCGTCA GTTGCCCTCCC

481 TTCTCCCAAT TCCTGCCGTT AGTCAACAAA ATTCCCTTTG AGCCTGTGCT TGAAGACTTT

541 GACTTCTATG GACATCAATT GGATGTTGAA TCAGCTGATC CTTTGAGCAT TTTAGGACAA

601 AACCCCTGGTT TCATGAGTCC GAGCTATGTC TTC

Gadus morhua (cod) FGF21 gene coding sequence (SEQ ID NO: 176) (Ensembl accession no. ENSGMOT00000014151, which is hereby incorporated by reference in its entirety)

1 CTCCCTCCCG CCACCCCTCTT CCACATCGGC CTCTCCTCT AC GTCCCCGA CTCCGGCCCC

61 CTGCTGTGGC TGGCGACCA GGTCAGGGAG AGACACCTCT ACACAGCAGA GAGCCACCGG

121 AGGGGGCTGT TCCTGGAGAT GAGCCCGGAC GGTCAGGTGA CAGGAAGTGC TGCTCAGACG

181 CGCCTCAGTG TTCTGGAGCT GAGGTGGTC AGAGCAGGAG ATACGGTCAT CAGAGCGCGC

241 CTCTCCTCTC TCTACCTGTG TGTGGACAGG GCAGGTCAAC TGACAGGACA GAGACAGTAC

301 ACAGAGTCCG ACTGCACCTT CAGAGAGGTC ATCCTTGAGG AC GGCTACAC CCACCTCCTG

361 TCCGTGCACC AC GGACTTCC TATTCGCTG GCGCCGAGAC ACTCCCCAGG GAGACAGGGG

421 CTGCGCTTCA GCAGGTTCTT CCCGCTGAGG AGCAGTGTG CAGAGGATAG GGTCGCCGAG

481 CCCCCAGACA GCCCACTGAA CCTGGACTCT GAAGACCCCC TGAGGATGGG TCTGGGTTCG

541 CTCCCTCAGCC CGGCCTTCTC CATG

TABLE 4-continued

Latimeria chalumnae (coelacanth) FGF21 gene coding sequence
(SEQ ID NO: 177) (Ensembl accession no. ENSLACT00000003815,
which is hereby incorporated by reference in its entirety)

1 ATGTTATGCC AGAGTTTGT GATATTAAGT CAGAAATTCA TTTTTGGCT CTTTTGACT

61 GGATTGGGC TAACAGGATT GGCTGGACA AGGCCCTTC AGGATTCAA TCCCACCTG

121 CAGTATTCCG ATTCCATCCG GCTCCGACAT CTGTACACTG CCAGTGAGAG TCGGCACCTT

181 CACCTACAAA TCAACTCGGA TGGACAGGTG GGAGGGACAA CCAAGCAAAG CCCTTACAGT

241 CTGTTGGAGA TGAAGGCGGT GAAGACAGGT TTTGTGGTCA TCAGGGCAA GAAAAGCGCC

301 CGTTACCTCT GTATGGAACG TAGTGGACGG CTCTATGGAT CGCTGCAGTA TACAGAAAAA

361 GACTGCACCT TCAAAGAGGT TGTGTTGGCA GATGGATACA ACCTGTATGT CTCAGAGGAA

421 CACCAGGCCA CAGTGACGCT GAGCCCCATG AGGGCGAGGA TAGCGCAAGG GAAAAGATC

481 CCACCCCTTT CCCATTCCT TCCAATGGTG AACAAAGGTGC CTGTGGAGGA TGTTGCCGCT

541 GAGATGGAGT TTGTCCAGGT GCTGCGGGAA ATGACGGCCG ACGTGGACTC TCCGGATCCC

601 TTTGGAATGA CCTGGGAAGA ATCGGTTCAC AGTCGGAGCT TTTTGCC

Tursiops truncatus (dolphin) FGF21 gene coding sequence
(SEQ ID NO: 178) (Ensembl accession no. ENSTTRT00000014561,
which is hereby incorporated by reference in its entirety)

1 ATGGGCTGGG ACAAGACCAA ACTCGAGCAC CTGGACTGT GGGTCCCTGT GCTAGCTGTC

61 CTGCTGGGAC CCTGCCAGGC ACATCCCATT CCTGACTCCA GCCCCCTCCT CCAATTGGG

121 GGCCAAGTCC GCCAGCGATA CCTCTACACG GATGACGCC AGGAGACGGA GGCCCACCTG

181 GAGATCAGGG CTGATGGCAC AGTGGTGGGG ACGGCCGCC GGAGCCCGA AGGAGTTAAA

241 ACATCCAGGT TCCTGTGCCA GGGGCCAGAG GGGAGGCTGT ATGGATCGCT CCACTTAAC

301 CCCCAGGCCT GCAGCTCCG GGAGCTGCTT CTTGAGGATG GATAAACGT TTACCAAGTCT

361 GAGGCTCTTG GCATTCCTC CCGCCTGCC CCGCACCGCT CCTCCAACGT GGACCTGGCC

421 CCCCGGGGAC CTGCTCGCTT CCTGCCGCTG CCAGGCTTCC TCCCGCCACC CCTGGAGCCT

481 CCAGGGATCT TGGCCCCGAA GCCTCCAAAC GTAGGTTCTT CGGACCCCTT GAGCATGGT

541 GGACCTTCAC ATGGCCGAAG CCCAGCTAC ACTTCCTGA

Mustela putorius furo (ferret) FGF21 gene coding sequence
(SEQ ID NO: 179) (Ensembl accession no. ENSMPUT00000003755,
which is hereby incorporated by reference in its entirety)

188 ATG GGCTGGGAAG AGGCAGGTC CGAGCACCTG GGGCTGTGGG TCCCTGTGCT

241 GGCGGTCCTT TTGCTGGAG CCTGCCAGGC ATACCTATT CCTGACTCCA GCCCCCTCCT

301 CCAATTGGGAA GGCAAGATTC GACAGCGGTA CCTCTACACA GACGACGCTC AGGAGACGGA

361 GGCCCACCTA GAGATCAGGG CTGATGGCAC GGTGGTGGGG GCTGCCGCC GGAGCCCGA

421 AAGTCTCTTG GAGCTGAAAG CCCTGAAGCC AGGGGTCTT CAGATCTTG GAGTAAAAC

481 ATCCAGGTTCTG CTGTGCCAGG GCCGAATGG GACACTGTAC GGATCGTTCC ACTTCGACCC

541 CGTAGCCTGC AGCTCCGGG AAGTGTCTT GGAAGATGGA TACAACATCT ACCACTCTGA

601 GACCCCTGGGCT CTCCCACTGC GCCTGCCCTTCC CCACAACTCC CCACACAGGG ACCTGGCGCC

661 CGGGGGGCCT GCCCGCTTCC TGCCCCTGCC AGGCCTGCTT CGGGCCACCC CGGAGTCCCG

721 GGGGATCCCA GCCCCGAGC CTCCCAACGT GGGCTCCTCA GACCCCTGA GCATGGTGGG

781 GCCTTTGCAG GGTCAAAGTC CCAGCTACAC TTCCTGA

Takifugu rubripes (fugu) FGF21 gene coding sequence (SEQ
ID NO: 180) (Ensembl accession no. ENSTRUT00000034076,
which is hereby incorporated by reference in its entirety)

1 TTTATTTATT TATTATTCAG AACTGACTT TTTTCCCTT CCAAATGGTT CAACTTTAT

61 CTCCCTGACT CCAACCCGCT CTTATCCTT GACAGTCATG GCAGAGGCAT CCACCTCTAC

TABLE 4-continued

121 ACAGATAATC AAAGGGGAGG GATGTATCTG CAGATGAGCA CAGATGGAAG CGTTTCCGGG
 181 AGTGATGTCC AGACGGCGAA CACTGTGCTG GAACTGAAGT CAGTCAGAAA CGGCCACGTC
 241 GTCATCCGAG GAAAATCGTC TTCTCTGTT CTCTGTATGG ACAGCAGAGG CCGTTTATGG
 301 GGGCAGAGGC ACCCCACTGA GGCGACTGC ACTTCAGGG AAGTGTGCT GGCAGATGGA
 361 TACACTCGCT TCCTGTCCT GCACAACGGA ACTCCTGTG CTCTGGCACC TAAACAATCT
 421 CCAGACCAGC ACACAGTCCC CTTCACTCGT TTCCTGCCGC TCAGGAATAC ACTGGCAGAG
 481 GAGAGCATGT CTGAACCAAC ATCAAACCAA CAGAGATATT TTAACATTGA CTCTGATGAT
 541 CTTCTGGAA TGGATTAAA TGCGATGGTC AGTCCTCAGT TTTCAGGGGA CAAGTGA

Dipodomys ordii (kangaroo rat) FGF21 gene coding sequence
 (SEQ ID NO: 181) (Ensembl accession no. ENSDORT00000001234,
 which is hereby incorporated by reference in its entirety)
 1 ATGGACCAGG CAAAGACCAG GGGTGGGGCC CGGGGGCTGG GGGGCCTGT GCTGGCTGTC

61 ATAATTCTGG GAGCATGCAA GGCACGGCCT ATCCCTGACT CCAGCCCCCT CCTCCAATT
 121 GGGGGTCAAG TTGGGCTTCG GCACCTCTAC ACAGATGACA CTCAGGAGAC GGAAGCCCAT
 181 CTGGAGATCA GGGCAGATGG CACGGTAGTG GGGACTGCC ACCGGAGCCC TGAAAGTCTC
 241 TTGGAGCTGA AAGCCTTGAA GCCAGGAGTC ATTCAAATCT TAGGGATCAA GACATCCAGA
 301 TTCTTATGCC AGAGACCAGA CGGGACACTG TATGGATCAC TCCACTTTGA CCCTGAGGTT
 361 TGCAGCTTCC AGGAGCTGCT TCTGGAAGAT GGATACAACA TTTACCGTTC TGAAGCCCTG
 421 GGTCTCCCCC TGCGCCCTGTC CCCAGATCCA GCACCCCTGGG GGCCAGCCCG CTCCCTGCC
 481 CTGCCTGGTG TGCCCCCGC ACCGCCGGAG CCCCCCGGGA TCCCTGGCTCC CGAACCCCC
 541 GATGTCGGCT CCTCCGACCC TCTGAGTATG GTGGGACTGT TGCAGGGCCG AAGCCCCAGC
 601 TATGCATCCT GA

Echinops telfairi (lesser hedgehog tenrec) FGF21 gene
 coding sequence (SEQ ID NO: 182) (Ensembl accession no.
 ENSETET00000010721, which is hereby incorporated by
 reference in its entirety)

1 ATGGGTTGCA CCAAATCTGG GTGGAAGTCC CGGGGACTGT GGGTCCCTGT GCTGGCCAGC
 61 CTTCTGCTGG GAGGCTGCGG AGCACACCCC ATCCCTGACT CCAGCCCCCT CCTCCAATT
 121 GGGGGCCAAG TCCGGCAGCG ATACCTCTAT ACGGATGACG CCCAGACCAC CGAGGCCAC
 181 CTGGAGATCA GAGCGGATGG CACAGTGGGG GGCGCTGCC ACCAGAGCCC AGAGAAGTTC
 241 CTGAGTCAAT GGCCTGAAAA GCCCCCTGAGA TCACTCCATT TCGACCCAGC CGCCTGCAGC
 301 TTCCGGGAGA AGCTCTAGA AGACGGATAC AACTTGTACC ACTCTGAGAC CCACGGCCTC
 361 CCCCTCCGCC TCCCCACCCCG TGGGGCGAC CCCTCTTCTC AGCCTGGGGC CCGCTCCCCA
 421 CCGCTGCCGG GCCAGCTCCC ACAACTCCAA GAGACGCCAG GGGTCCCTCGC CCCCCGAACCC
 481 CCCGACGTGG GCTCTTCAGA CCCCCCTGAGC ATGGTGGGGC CTTGGCGAGG GCAAAGTCCC
 541 AGTTATGCCT CCTGA

Macaca mulatta (rhesus monkey) FGF21 gene coding sequence
 (SEQ ID NO: 183) (Ensembl accession no. ENSMMUT00000038440,
 which is hereby incorporated by reference in its entirety)
 1 ATGGACTCGG ACGAGACCGG GTTCGAGCAC TCAGGACTGT GGGTCCCTGT GCTGGCTGGT

61 CTTCTGCTGG GAGCCTGCCA GGCACACCCC ATCCCTGACT CCAGTCCTCT CCTGCAATT
 121 GGGGGCCAAG TCCGGCAACG GTACCTCTAC ACAGATGATG CCCAGCAGAC AGAAGCCCAC
 181 CTGGAGATCA GGGAGGATGG GACAGTGGGG GGCGCTGCTC ACCAGAGCCC CGAAAGTGAG
 241 TGTGGGCCAG AGCCTGGGTG TGAGGGAGGA GGGGCTGTGG GAGGTGCTGA GGGACCTGGA
 301 CTCCTGGGTC TGAGGGAGGC AGGGCTGGGG CCTGGATCCT GGCTCCACTT TGACCCCTGAG

TABLE 4-continued

361 GCCTGCAGCT TCCGGGAGCT GCTTCTTGAG AACGGATACA ATGTTTACCA GTCCGAGGCC
 421 CACGGCCTCC CACTGCACCT GCCGGAAAC AAGTCCCCAC ACCGGGACCC TGCACTCCAA
 481 GGACCAGCTC GCTTCCTGCC ACTACCAGGC CTGCCCGGG CACCCCGGA GCCGCCAGGA
 541 ATCCTCGCCC CCCAGCCCC CGATGTGGGC TCCTCGGACC CTCTGAGCAT GGTGGGACCT
 601 TCCCAGGCC GAAGCCCCAG CTATGCTTCC TGA

Microcebus murinus (mouse lemur) FGF21 gene coding sequence
 (SEQ ID NO: 184) (Ensembl accession no. ENSMICT00000013258,
 which is hereby incorporated by reference in its entirety)
 1 ATGGGCTGGG ACGAGGCCGG CGCCGGGTT GAGCACCCAG GACTGTGGTT TCCCATGCTG
 61 GGTGTCTGC TGCTGGGAGC CTGCCAGGCG TACCCCATCC CTGACTCCAG CCCCTCCTC
 121 CAATTGCGC GCCAAGTCCG GCAGCGGCAC CTCTACACAG ACGATATCCA GGAGACAGAA
 181 GCCCACCTGG AGATCAGGGC GGACGGCACA GTGGTGGGG CCGCCCGACA GAGCCCTGAG
 241 TTGGAGCTGA AAGCCTAAA GCCAGGGGTC ATTCAAATCT TGGGAGTCAA GACCTCCAGG
 301 TTCCTGTGCC AGAGGCCAGA CGGGGCCCTG TACGGATCGC TCCACTTTGA CCCCGAGTGC
 361 AGCTTCCGGG AGCTGCTTCT TGAGGATGGA TACAACGTCT ACTGTCCCTA CCTCCCGCTG
 421 CACCTGTCCC CACGCATCGA ACTGGCCGGA TCACGCTCTG CGCTGCCACT GCCCCCAGCA
 481 CCTGAACGCA GGATTTGGC CCCGGAGGCC CCGGATGGCT CCTCGGACCC TCTGAGCATG
 541 GTGGGGCCTT CGCAGGCCCG AAGTCCCAGC TATGCTTCTT GA

Ochotona princeps (pika) FGF21 gene coding sequence (SEQ
 ID NO: 185) (Ensembl accession no. ENSOPRT00000007373,
 which is hereby incorporated by reference in its entirety)
 1 AAAGACATGG ACGGGCTCCA GCCTCCGGGG CTGCGGGTT CTGTGCTGGC TGCCCTGCTT
 61 TTGGGAGTTG GCCAGGGCACG CCCCATCCCT GATTCTAGCC CTCTCTCCA ATTGGGGGC
 121 CAGGTCCGGC AGAGGCACCT CTACACGGAT GACGCCAGG AATCGGAAGT ACACCTGGAG
 181 ATCCGGGCAG ACGGCACCGT GGCAGGGACT GCCCGCCGGA GCCCTGAAAG TCTCTTAGAA
 241 ATGAAAGCGT TGAAGCCAGG CGTCATTCACT ATCCTGGGG TCCACACATC CAGGTTCCCTG
 301 TGCCAGAGAC CAGACGGGAC GCTGTACGGC TCGCTCCACT TCGACCACAA GCCCTGCAGC
 361 TTCCGGGAGC AGCTGCTGGA GGATGGGTAC AACGTGTACC ACTCAGAGAC ACACGGCCTC
 421 CCGCTGCGCC TGTCTCCAGA CCGAGCCCC CGGGGCCAG CCCGCTTCCCT GCCACTGCCA
 481 GCCCCTCCTC CTGACCTCCT GGTGCCACCC CTGCCACCGG ACGTCCCTAGC CCCTGAGGCC
 541 CCCGACGTGG ACTCCCCAGA CCCCCCTGAGC ATGGTGGGGC CCTTGCAGGG CCAAAGCCCC
 601 AGCTACACTT CCTGA

Xiphophorus maculatus (platyfish) FGF21 gene coding sequence
 (SEQ ID NO: 186) (Ensembl accession no. ENSXMAT00000001579,
 which is hereby incorporated by reference in its entirety)
 1 TGGCCGTTC CCTTCCTTT CTTAACCTCTC TCTCTCTTCTC CTCGTTTAC
 61 ATCCCCAGAAT CCAACCCAAT CTTTGCTTC AGGAATCAGC TCAGAGAGGT GCATCTCTAC
 121 ACAGAAAATC ACAGACGGGG TTTGTATGTG GAGATACATC TGGATGGGAG AGTGACTGGA
 181 AGTGATGCTC AGAGTCCTTA TAGTGTGTTG CAGATAAAAGT CTGTTAAACC GGGTCATGTG
 241 GTCATAAAGG GACAGACATC GTCCCTGTTC CTCTGCATGG ACGACTCCGG GAATCTAAGA
 301 GGACAGACAA CCTATGACGA GGCTGACTGC TCCTTCAGGG AACTGCTGCT GGCGATGGC
 361 TACACCCGTT CCCTGAACTC ACAACATGGC GTTCCTTTAT CACTGGCATC CAGAAACTCT
 421 CCAGATCGAC ACTCCGTTCC TTTACAAGA TTTTACCTC TCAGGAATAC TTTAACGGTT

TABLE 4-continued

481 TCAGAAGAAC CAACAAAAAC TCAGAGGGAC TTCAACCTGG ACTCGGACGA CCTTCTCGGG

541 ATGGGA

Gasterosteus aculeatus (stickleback) FGF21 gene coding sequence (SEQ ID NO: 187) (Ensembl accession no. ENSGACT00000010725, which is hereby incorporated by reference in its entirety)

1 TCTCTCTCC TCATGGTCCC ACTTCCTTC TGTTCATCCT TTTATCTCAC TGACTCCAGC

61 CCACCTTAC CTTCAATAA TCAAGTCAA GAGGTGCACC TCTACACAGC AGAGAATCAC

121 AGAAGAGCGA TGTACCTGCA GATCGCTCTG GACGGGACCG TGTCGGGAAG CGACGCTCGG

181 TCCACTTACA GTGTGCTGCA GCTGAAATCT ATCCAGCCGG GCCACGTGGT CATCAGAGGG

241 AAGGCCTCCT CCATGTTCT CTGCGTGGAC AGCGGGGGCC GTTGAGAGG ACAGGGGCCG

301 TACTCAGAGG CCGACTGCAG CTTCAGGGAG CTGCTGCTGG GGGATGGCTA CACCCGGTTC

361 CTGTCCTCGC AGCACGGGTC CCCGCTGTCT CTGGCGTGA GGCTTCCCC GGATCCCAAC

421 TCGGTGCCCT TCACTCGATT CCTACCCATC CGGACCGCCC CCGAGGCTGA GAGCGTGATC

481 GAAGAGGCCAC CGAGCAATCA GAGATACGTC AACGTGGACT CCGAGGATCT TCTTGGAATG

541 GGCCCTGAACA CTGTGGTCAG TCCTCAGTTC TCGGCG

Sarcophilus harrisii (Tasmanian devil) FGF21 gene coding sequence (SEQ ID NO: 188) (Ensembl accession no.

ENSSHAT00000006017, which is hereby incorporated by reference in its entirety) (1-209, excluding 1-2 and 173-209)

132 GTGTCTGCC ATGGGCCTGA GGGAGCGAGC TCCCAGGTAC CTGGCCCCGC

181 TGCTGTCCTT GCTCTTGGCC TGCAGGGCCT CGGGTCACCC CCTCCCGGAT TCCAGCCCCA

241 TGCTCCTGTT TGGGGGGCAG GTCCGCCTCC GGACACCTCTA CACGGATGTG GGCCAGGAGG

301 CCGAGGGCCA CGTGGAACTG GCGTCCGACG GCACAGTCGG GGCGGCAGCG CGGAGGAGTC

361 CCAACAGTCT CCTGGAGCTG AAGGCTGTGA AGCCGGGCAT CGTCCGAATC CTGGCCGTCC

421 ACAGCTCTCG GTTTCTGTGT ATGAGGCCA ACGGGGAGCT GTACGGAGCG ATACACTACG

481 ACCCTTCCGC CTGCAACTTT CGGGAGCGCC TGCTGGGGGA CGGCTACAAC GTGTACGAGT

541 CGGAGGCTCA CGGGAGGACC CTCCGCCTGC CCCCCAAGGC CGCACCGGGGAC CCCGCCGGAC

601 CTTCTCGCTT CCTGCCGCTC CCCGGC

Macropus eugenii (wallaby) FGF21 gene coding sequence

(SEQ ID NO: 189) (Ensembl accession no. ENSMBOU00000015309, which is hereby incorporated by reference in its entirety)

1 ACAGAGGAGC CTTCTACTGG GTCCAGGGCAC CTGGGACAAT GGGCTCCCGG GCTGCCTGGT

61 CCTCTGCTGT CCTTGCTCCT GGCCTACAGG GGCTGGGCT CCCCCATCCC TGATTCCAGC

121 CCCATGCTCC TGTTGGTGG CCAGGTCCGC CTCCGACACC TGTACACAGA TGATGCCAG

181 GACACGGAGG CCCATGTGGA GCTGGGGCCA GATGGAGTGG TTCGAGCTGT GGCTGAGAGG

241 AGCCCCAACAA GTCTCTGGA ACTGAAGGG GTGAAGCTG GAGTCATCCG AATCCTCGCT

301 GTCCAGAGCT CTCGGTTCT GTGTATGAGG CCCAACGGGG AACTGTATGG AGCGGTACAC

361 TATGACCTT CTGCCTGCAA CTTTGGGAA CATCTGCTGG GGGATGGTTA TAATGTGTAT

421 GAATCAGAGA CTCACAGAAC GACCCCTCCGT CTGTCCTCAT CCCTGGGTCA GGCTGGCCCC

481 TCTCGCTTCC TGCCACTTCC AGGCGACTGG CTGCCCGGCC CTGATCCACC TTGGGCACAG

541 GGCCCTGAGC CCCCCAGACGT GGGCTCTGCA GACCCCTGA GCATGGTGGG GGCCGTGCAG

601 GGCCCTCAGCC CCAGCTACTC CTCCTGA

TABLE 4-continued

Xenopus tropicalis (Western clawed frog) FGF21 gene coding sequence (SEQ ID NO: 190) (Ensembl accession no. ENSXETT0000009917, which is hereby incorporated by reference in its entirety) (1-209, excluding 170-209)

1 AGAGGGGTA GGACCAAAAA AAAGACGTTA CTCAGGAAT GGCTTGCGCT TTTAGCCATT

61 ATGTTGAGTA GGTCAAGGTT TTCTTAGCA AATCCTATCC AGAATTCGAA CCCAATCTTA

121 TCCAACGACA ACCAAGTACG GACTCAGTAT TTATACACAG ATAACAATAA CATGCACCTG

181 TATCTTCAGA TCACCCACAA TGAGTAGTA ACTGGTACCG AAGAAAAGAA TGACTATGGT

241 GTGCTGGAAA TAAAGGCAGT AAAAGCTGGG GTTGTAGTTA TAAAAGGAAT TCGAAGCAAT

301 CTCTACCTAT GCATGGATT TAGACACCAA TTGTATGCGT CGGCATATGA TAAAGATGAC

361 TGCCATTTC ATGAAAAGAT CACACCAGAT AATTACAACA TGTATAGCTC AGAGAACAT

421 TCAGAATACG TGTCTTAGC TCCATAAAAA GGAAGCCAGA TGGCTCGTTT TCTACCTATA

Danio rerio (zebrafish) FGF21 gene coding sequence (SEQ ID NO: 191) (Ensembl accession no. ENSDART00000103511, which is hereby incorporated by reference in its entirety)

30 A TGCTTCTGC CTGCTTTTT ATATTTTTG

61 CTCTTTTC TCATCTCGG TGGTGTATGT ATGTTCTGC ACAGAACGTG CTTCTGCAGT

121 TTGGCACACA AGTCAGGGAA CGCCTGCTTT ACACAGATGG GTTGTTCCTT GAAATGAATC

181 CAGATGGCTC CGTCAAAGGC TCTCTGAAA AGAATCTAAA TTGTGTGCTG GAGCTGCGTT

241 CAGTCAAAGC GGGTGAAACC GTCATCCAGA GTGCAGCTAC ATCTCTCTAC CTCTGCGTCG

301 ATGATCAAGA CAAGCTGAAA GGACAGCATC ATTACTCTGC ACTAGACTGC ACCTTCAGG

361 AATTGCTACT GGATGGATAT TCGTTTTCC TTTCTCCACA CACTAATCTT CCCGTATCGC

421 TCCTCTCGAA ACGTCAGAAA CACGGCAATC CTCTTCTCG CTTCCTCCCT GTTACGAG

481 CAGAGGACAG CGGGACACAG GAGGTGAAAC AGTATATTCA GGATATAAAC CTGGACTCTG

541 ACGACCCACT AGGAATGGGA CATCGGTAC ACTTACAGAC CGTCTTCAGT CCCAGTCTGC

572 ATACTAAAAA ATGA

Bos grunniens mutus (yak) FGF21 gene coding sequence (SEQ ID NO: 192) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org/))

1 ATGGGCTGGG ATGAAGCGAA ATTTAACAT CTGGGCCTGT GGGTGCCGGT GCTGGCGGT

61 CTGCTGCTGG GCACCTGCCG CGCGCATCCG ATTCCGGATA GCAGCCCGCT GCTGCAGTTT

121 GGCGGCCAGG TGCGCCAGCG CTATCTGTAT ACCGATGATG CGCAGGAAAC CGAACGCGAT

181 CTGGAAATTG GCGCGGATGG CACCGTGGTG GGCGCGGCGC GCCAGAGCCC GGAAAGCCTG

241 CTGGAACCTGA AAGCGCTGAA ACCGGCGTG ATTCAAGATC TGGGCGTGAA AACCAAGCCG

301 TTTCTGTGCC AGGGCCCGGA TGGCAAATG TATGGCAGCC TGCACTTTGA TCCGAAAGCG

361 TGCAGCTTTC GCGAACCTGCT GCTGAAAGAT GGCTATAACG TGTATCAGAG CGAAACCCCTG

421 GGCCCTGCCG TGCGCCCTGCC GCCGCGAGCGC AGCAGCAACC GCGATCCGGC GCCCGCGCGC

481 CGGGCGCGCT TTCTGCCGCT GCCGGGCCTG CGGGCGGAAC CGCCGGATCC GCCGGGCATT

541 CTGGCGCGGG AACCGCCGGA TGTGGGCAGC AGCGATCCGC TGAGCATGGT GGGCCCGAGC

601 TATGGCCGCA GCCCGAGCTA TACCAAGCTAA

Saimiri boliviensis boliviensis (Bolivian squirrel monkey) FGF21 gene coding sequence (SEQ ID NO: 193) (GenBank accession no. XM_003940326, which is hereby incorporated by reference in its entirety)

163 atgggctc ggaggaggtc

181 GCGTTGGAGC GCCCTGCACT GTGGGTCTCT GTGTTGGCTG GTCTCCTGCT GGGAACCTGC

241 CAGGCATACC CCATCCCTGA CTCTAGTCCC CTCCCTGCAAT TTGGAGGCCA AGTCCGGCAG

TABLE 4-continued

301 CGGTACCTCT ACACAGATGA CGCTCAGCAG ACAGAAGCCC ACCTGGAGAT CAGGGAAGAT
 361 GGCACGGTGG CGGGGCTGC CCACCAGAC CCCGAAAGTC TCTTGAGCT GAAAGCCTTA
 421 AAGCCAGGGG TTATTCAAAT CTTGGGAGTC AAGACCTCCA GGTTCCCTGTG CCAGAGGCCG
 481 GACGGGGCCC TGTACGGATC GCTCTACTTT GACCCCGAGG CCTGCAGCTT CCGGGAGCTG
 541 CTTCTTGAGG ACGGATACAA TGTGTACCAAG TCCGTGGCCC ACAGCCTCCC GCTGCACCTG
 601 CCAGGGGGCA GGTCCCCACC CTGGGACCT GCACCTCGAG GACCAGCTCG CTTCCGTGCG
 661 CTACCAGGCC TGCCCCCGA ACCCCCCGAG GCGCCAGGAA TCCGTGGCCC CGAGCCCCCC
 721 GATGTGGGCT CCTCAGACCC TCTGAGCATG GTGGGGCCTT CCCAAGGCCA AAGCCCCAGC
 781 TACACTTCCT GA

Callithrix jacchus (white-tufted-ear marmoset) FGF21 gene coding sequence (SEQ ID NO: 194) (GenBank accession no. XM_003735621, which is hereby incorporated by reference in its entirety)

1 ATGGGCTCGG AGGAGGTCGG GTTGGAGCAC CCTGCACTGT GGGTTCTGT GCTGGCTGGT
 61 CTCCTGCTGG GAACCTGCCA GGCGCACCCCC ATCCCTGACT CCAGTCCCCCT CCTGCAATT
 121 GGAGGCCAAG TCCGGCAGCG GTACCTCTAC ACAGATGACG CCCAGCAGAA AGAACGCCAC
 181 CTGGAGATCN AGGAAGATGG CACAGTGGCC GGGCTGCCA CCAAAGTCCC GAAAGTGAGT
 241 CTCTTGAGCG TGAAAGCCTT AAAGCCAGGG GTTATTCAAAT TCTTGGGAGT CAAGACATCC
 301 AGGTTCTGT GCCAGAGGCC AGACGGGGCG CTGTATGGAT CGCTCCACTT TGACCCCGAG
 361 GCCTGCAGCT TCCGGGAGCT GCTCTTGAG GACGGATACA ATGTGTACCA GTCTGTGGCC
 421 CACGGCCTCC CGCTGCACCT GCCAGAGAGC AGGTCACCAC CCCGGGACCC TGCACCCCGA
 481 GGACCAGCTC GCTTCTGAGC ACTACCAGGC CTGCCCCCTG AACCCCCAGA GCCGCCAGGA
 541 ATCCTGGCCC CTGAGCCCC CGACGTGGGC TCCTCAGACC CTCTGAGCAT GGTGGGGCCT
 601 TCCCAAGGCC AAAGCCCCAG CTACGCTTCC TGA

Tupaia chinensis (Chinese tree shrew) FGF21 gene coding sequence (SEQ ID NO: 195) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org))

1 ATGGGCTGGG ATAAAGCGCG CTTTGAACAT CTGGGCGCGT GGGCGCCGGT GCTGGCGGT
 61 CTGCTGCTGG GCGCGTGCCTA GGCATCG ATTCCGGATA GCAGCCCGCT GCTGCAGTT
 121 GGCGGCCAGG TGCGCCAGCG CTATCTGTAT ACCGATGATA CCCAGGATAC CGAACGCGCAT
 181 CTGGAAATTG GCGCGGATGG CACCGTGGTG GGCGCGCGC ATCAGAGCCC GGAAAGCCTG
 241 CTGGAACGTAA AAGCGCTGAA ACCGGCGTG ATTCAAGATTC TGGGCGTGAA AACCAAGCCG
 301 TTTCTGTGCC AGCGCCCGGA TGGCGCGCTG TATGGCAGCC TGCACTTTGA TCCGGAAGCG
 361 TGCAGCTTTC GCGAACATGCT GCTGAAAGAT GGCTATAACA TTTATCAGAG CGAACGGCGC
 421 GGCGCTGCCG TGCGCCCTGCC GCCGCATGAT AGCCCGCATC GCGATCGCAC CCCGCAGGGC
 481 CGGGCGCGCT TTCTGCCGCT GCCGGGCCTG CGCGTGGTGC CGCCGGAACG GCCGGCGTG
 541 CTGGCGCTGG AACCGCCGGA TGTGGGCAGC AGCGATCCGC TGAGCATGAT GGGCCCGAGC
 601 CAGGGCCAGA GCCCGAGCTA TGCGAGCTAA

Papio anubis (olive baboon) FGF21 gene coding sequence (SEQ ID NO: 196) (GenBank accession no. XM_003915851, which is hereby incorporated by reference in its entirety)

1 ATGGACTCGG ACGAGACCGG GTTCGAGCAC TCAGGACTGT GGGTTCTGT GCTGGCTGGT
 61 CTTCTGCTGG GAGCCTGCCA GGACACCCCCC ATCCCTGACT CCAGTCCCT CCTGCAATT
 121 GGGGGCCAAG TCCGGCAACG GTACCTCTAC ACAGATGATG CCCAGCAGAC AGAACGCCAC

TABLE 4-continued

181 CTGGAGATCA GGGAGGATGG GACAGTGGGG GGCGCTGCTC ACCAGAGCCC CGAAAGTAAG

241 TGTGGGCCAG AGCCTGGTC TGAGGGAGGA GGGGCTCTCC ACTTTGACCC TGAGGCCTGC

301 AGCTTCCCGCG AGCTGCTTCT TGAGAACGGA TACAATGTTT ACCAGTCCGA GGCCCACGGC

361 CTCCCCACTGC ACCTGCCGGG AAACAAGTCC CCACACCAGG ACCCTGCATC CCGAGGACCA

421 GCTCGCTTCC TGCCACTACC AGGCTGCC CCCGCACCCC CAGAGCCACC AGGAATCCTC

481 GCCCCCCAGC CCCCCGATGT GGGCTCCTCG GACCTCTGA GCATGGTGGG ACCTTCCCAG

541 GCCCGAAGCC CTAGCTACGC TTCCTGA

Pteropus alecto (black flying fox) FGF21 gene coding sequence (SEQ ID NO: 197) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org))

1 ATGGGCTGGG GCAAAGCGCG CCTGCAGCAT CGGGGCCTGT GGGGCCGGT GCTGGCGGTG

61 CTGCTGGCG CGTGCCAGGC GCATCCGATT CTGGATAGCA GCCCGCTGTT TCAGTTGGC

121 AGCCAGGTGC GCGCCGCTA TCTGTATACC GATGATGCGC AGGATACCGA AGCGCATCTG

181 GAAATTCCCG CGGATGGCAC CGTGGCGGGC GCGGCGCGCC GCAGCCCGGA AAGCCTGCTG

241 GAACTGAAAG CGCTGAAACC GGGCGTGATT CAGGTGCTGG GCGTGAAAAC CAGCCGCTTT

301 CTGTGCCAGC GCCCCGATGG CACCCGTAT GGCAGCCTGC ATTTGATCC GGCGCGGTGC

361 AGCTTCCGCG AACTGCTGCT GAAAGATGGC TATAACGTGT ATCAGAGCGA AGCGCTGGCG

421 CGCCCGCTGC GCCTGCCGCC GTATAGCAGC CCGAGCAGCG ATCCGGCGCG CC CGCGGCCCG

481 GCGCGCTTC TGCCGCTGCC GGGCCCGCC CGGAAACCGC CGCAGCCGCC GGGCCGCCCTG

541 GCGCCGGAAC CGCCGGATGT GGGCAGCAGC GATCCGCTGA GCATGGTGTG GCCGAGCCGC

601 GGCGCGCAGCC CGAGCTATAC CAGCTAA

Heterocephalus glaber (naked mole-rat) FGF21 gene coding sequence (SEQ ID NO: 198) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org))

1 ATGGATTGGG CGCGCGCGGA AAGCGAACGC CGGGGCCTGT GGGTGCCGGC GGTGCTGGCG

61 GTGCTGCTGC TGGGCGCGTG CCAGCGCAT CCGATTCCGG ATAGCAGCCC GCTGCTGCAG

121 TTTGGCGGCC AGGTGCGCCA GCGCCATCTG TATACCGATG ATGCGCAGGA TACCGAAGTG

181 CATCTGGAAA TTGCGCGCGA TGGCAGCGTG GGCAGCGCGG CGCATCGCAG CCCGAAAGC

241 CTGCTGAAAC TGAAAGCGCT GAAACCGGGC GTGATTAGA TTCTGGCGT GCGCACCAGC

301 CGCTTTCTGT GCCAGCGCCC GGATGGCACC CTGTATGGCA GCCTGCATTT TGATCCGGAA

361 GCGTGCAGCT TTCGCGAACT GCTGCTGGCG GATGGCTATA ACATTTATCA GAGCGAAGCG

421 TATGGCCTGC CGCTGCGCAT GCTGCCGAGC GATAGCGCGA GCGCGATCC GGTGCCGCCG

481 GGCCCGGCCG GCTTCTGCC GCTGCCGGGC CTGCATCCGC CGCCGCTGGA ACCGCCGGGC

541 ATGCTGCCGC CGGAACCGCC GGATGTGGGC AGCAGCGATC CGCTGAGCAT GGTGGCCCG

601 CTGCAGGGCC GCAGCCCGAG CTATGCGTTT TAA

Cricetulus griseus (Chinese hamster) FGF21 gene coding sequence (SEQ ID NO: 199) (GenBank accession no. XM_003508678, which is hereby incorporated by reference in its entirety)

1 ATGGACTGGG TGAAATCTGG AGTTGGGGTC CGGGACTGT GGGTCCCTCT GCTGCCTATC

61 TTCCCTGCTGG GGGTCTCCCA GGCACACCCC ATCCCTGACT CCAGCCCCCT CCTCCAGTTT

121 GGGGGTCAAG TCCGGCACAG GCACCTCTAC ACAGATGACA ACCAGGAAAC TGAAGTCCAC

181 CTGGAGATTA GGCAGGATGG CACGGTGATA GGGAACACAC ACCCGAGCCC AGAAAGTCTC

241 CTGGAGCTCA AAGCCTGAA GCCAGAGGTC ATCCCAGTGC TGGGTGTCAA GGCCTCCAGG

TABLE 4-continued

301 TTTCTTGC AACAACCAGA CGGAACCCCTG TATGGATCGC CTCACTTGA TCCTGAGGCC
 361 TGCAGTTCA GGGAGCTCTT GCTTGAGGAT GGATACAATG TGTACAAATC TGAAGTCCAT
 421 GGCCTGCCCT TGCGCCTGCC CCAGAGGGAC TCTCCAAACC AGGCCCCAGC ATCCTGGGA
 481 CCTGTGCCCT CCCTGCCAGT GCCAGGACTG CTCCACCAGC CCCAGGAGCT ACCAGGGTTC
 541 CTGGCCCCAG AACCTCCAGA TGTGGGCTCC TCTGACCCAC TGAGCATGGT GGGACCTTG
 601 CAGGGCCGAA GCCCCAGCTA TGCTTCCCTGA

Ovis aries (sheep) FGF21 gene coding sequence (SEQ ID NO: 200) (GenBank accession no. XM_004015796, which is hereby incorporated by reference in its entirety)

1 ATGGGCTGG ACAGGAGCAA GTTCAAGCAC TTGGGACTGT GGGTCCCTGT GCTGGCTGTC
 61 CTCCTGCTAG GAACCTGCCG GGCGCATCCA ATTCCAGACT CCAGCCCCCT CCTCCAGTTT
 121 GGGGGCCAAAG TCCGCCAGCG GTACCTCTAC ACGGATGATG CCCAGGAGAC AGAGGCCAC
 181 CTGGAGATCA GGGCGATGG CACAGTGGTG GGGCGGGCC GCGAGAGTCC CGAAAGTCTC
 241 TTGGAGCTGA AAGCCTGAA GCCAGGAGTC ATTCAAGATCT TTGGAGTTAA AACATCCAGG
 301 TTCCTGTGCC AGGGGCCAGA TGGGAAGCTG TATGGATCGC TGCACCTTGA CCCAAAGCC
 361 TGCAGCTTCC GGGAGCTGCT TCTTGAAGAT GGGTACAATG TCTACCAGTC GGAGACCCCTG
 421 GGCCTTCCAC TCCGCCCTGCC GCCGCCAGGGC TCATCCAACC GGGACCCGGC CCCGCGGGGA
 481 CCTCCGAAGC CCCAGCTACA CTTCTTGAAG ACGTCGCTG TGCAGTACTG GCCACGTTAT
 541 GAGAAGGTCC CAGCTTTCT GCACCCCTTC CCCGGCTGA

Pan paniscus (pygmy chimpanzee) FGF21 gene coding sequence (SEQ ID NO: 201) (GenBank accession no. XM_003814115, which is hereby incorporated by reference in its entirety)

(1-209, excluding 117-194 and 202-209)

573	ATGGACTC GGACGGAGACC GGTTGGAGC
601	ACTCAGGACT GTGGGTTCT GTGCTGGCTG GTCTTCTGTG GGGAGCCTGC CAGGCACACC
661	CCATCCCTGA CTCCAGTCCT CTCCTGCAAT TCGGGGGCCA AGTCCGGCAG CGGTACCTCT
721	ACACAGATGA TGCCCAGCAG ACAGAAGGCC ACCTGGAGAT CAGGGAGGAT GGGACGGTGG
781	GGGGCGCTGC TGACCAAGAGC CCCGAAAGTC TCCTGCAGCT GAAAGCCTTG AAGCCGGGAG
841	TTATTCAAAT CTTGGGAGTC AAGACATCCA GGTTCCCTGTG CCAGAGGCCA GATGGGGCCC
901	TGTATGGATC GGTGAGTTTC ----- ----- ----- ----- ----- ----- -----
921	----- ----- CAG ----- ----- ----- ----- ----- ----- ----- -----
924	----- ----- GAC CCTCCT ----- ----- CA CCACCCACCA ----- T
946	GCTCC----- TCCTAT ATGTCGCCCTCACAG----- CCTGGG

Macaca fascicularis (crab-eating macaque) FGF21 gene coding sequence (SEQ ID NO: 202) (generated using SMS Reverse Translate tool on the ExPASy Bioinformatics Resource website (www.expasy.org)) (1-209, excluding 117-209)

1	ATGGATAGCG ATGAAACCGG CTTTGAACAT AGCGGCCCTGT GGGTGGCGGT GCTGGGGGC
61	CTGCTGCTGG GCGCGTGCCA GGCGCATCCG ATTCCGGATA GCAGCCCGCT GCTGCAGTT
121	GGCGGCCAGG TGCGCCAGCG CTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT
181	CTGGAAATTC GCGAAGATGG CACCGTGGGC GGCGCGGGCGC ATCAGAGCCC GGAAAGCCTG
241	CTGCAGCTGA AAGCGCTGAA ACCGGCGTG ATTCAAGATTC TGGGCCTGAA AACCAAGCCG
301	TTTCTGTGCC AGAAACCGGA TGGCGCGCTG TATGGCAGCG TGAGCTTTA A

TABLE 4-continued

Mesocricetus auratus (golden hamster) FGF21 gene coding sequence (SEQ ID NO: 203) (GenBank accession no. EU497769, which is hereby incorporated by reference in its entirety) (1-209, excluding 1-89 and 194-209)

1 GGTCATCCAA ATCCTGGGTG TCAAGGCTGC TAGGTTTCT TGCCAGCAAC CAGACGGAAG

61 CCTGTACCGA TCGCCTCACT TCGATCCCGA GGCCCTGCAGT TTCCGGGAGC TCCTGCTTGA

121 GGATGGATAC AATGTGTACC AGTCGGAAGC CCACGGCCTG CCCCTGCCGC TGCCCCAGAG

181 GGACGCTCCG AGCCAGCCCC CAGCATCCTG GGGACCGGTG CGCTTCCTGC CAGTGCCCAG

241 ACTGTTCCAG CCGCCCCACG ACCTCCCAGG GCGCCCGGCC CCAGAGCCTC CGGACGTGGG

301 CTCCTCCGAC CCAC

Nile tilapia FGF21 gene coding sequence (SEQ ID NO: 204) (GenBank accession no. XM_003438468, which is hereby incorporated by reference in its entirety) (1-209, excluding 1-58)

1 ATGTATTTGC AGATGAACAT GGATGGGAGA GTCACAGGAA GTGATGCTCA GACACCTTAC

61 AGTTTGATGC AGCTGAAATC AGTAAACCA GGCCATGTAA TCATTAAGG ACCATCATCA

121 TCTCTTTTC TCTGTGTGGA CAGCGAAGGC AATCTGAGAG GGCAGAGTCA CTACTCAGAA

181 ACCAGCTGCA CCTTCAGAGA AATGCTGCTG GCTGACGGAT ACACCCGTTT CATTTCCTCA

241 CAATATGGAT TTCCCATGTC ACTGGCATCA AGACATTCCC CAGATCGACA CGCGCTTCCC

301 TTACGCGGT TCCTACCACT GAGGAATAAC TTGAAAACGG ATAGCGTATC AGAGCAGCTG

361 CCAAACAATC AGAGACTCTT CAACGTGGAC TCTGATGACC TTCTTGAAT GGGTCTAAAT

421 TCTATGGCA GTCCTCAGTT TTCTATGGAC AAATAA

In one embodiment, the chimeric protein of the present invention comprises the amino acid sequence of SEQ ID NO: 205, SEQ ID NO: 206, SEQ ID NO: 207, SEQ ID NO: 35, 208, SEQ ID NO: 209, or SEQ ID NO: 210, as shown in Table 5.

TABLE 5

Description of Chimeric Protein Sequence

Amino acid sequence of a FGF21/19 chimera composed of residues H29 to V197 of human FGF21 and residues T204 to K216 of human FGF19 (bold)	SEQ ID NO: 205 HP IPDSSPLQF GGQVRQRQLY TDDAQQTTEAH LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVTGL EAVRSPSFEK
Amino acid sequence of a FGF21/19 chimera composed of residues H29 to S190 of human FGF21 and residues M197 to K216 of human FGF19 (bold)	SEQ ID NO: 206 HP IPDSSPLQF GGQVRQRQLY TDDAQQTTEAH LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS MDPFGLVTGL EAVRSPSFEK
Amino acid sequence of a FGF21/19 chimera composed of the β-trefoil core domain of human FGF21 (residues H29 to L167) and the C-terminal tail of human FGF19 (residues L169 to K216) (bold)	SEQ ID NO: 207 HP IPDSSPLQF GGQVRQRQLY TDDAQQTTEAH LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL VPEEEPEDLRG HLESDMFSSP LETDSMDPFG LVTGLEAVRS PSFEK
Amino acid sequence of a FGF21/19 chimera composed of residues H29 to V197 of human FGF21 harboring Q104M mutation and residues T204 to	SEQ ID NO: 208 HP IPDSSPLQF GGQVRQRQLY TDDAQQTTEAH LEIREDTGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCMRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK

TABLE 5-continued

Description of Chimeric Protein Sequence	
K216 of human FGF19 (bold)	SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVTG L EAVRSPS F EK
Amino acid sequence of a FGF21/19 chimera composed of residues H29 to S190 of human FGF21 harboring Q104M mutation and residues M197 to K216 of human FGF19 (bold)	SEQ ID NO: 209 HP IPDSSPLLQF GGQVRQRYLY TDDAQQTTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCMRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS MDPFGLVTG L EAVRSPS F EK
Amino acid sequence of a FGF21/19 chimera composed of the β -trefoil core domain of human FGF21 (residues H29 to L167) harboring Q104M mutation and the C-terminal tail of human FGF19 (residues L169 to K216) (bold)	SEQ ID NO: 210 HP IPDSSPLLQF GGQVRQRYLY TDDAQQTTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCMRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPL LPM VP <small>E</small> PEDLRG HLESDMFSSP LETDSMDPFG LVTGLEAVRS PSFEK

In one embodiment of the present invention, the chimeric protein may include one or more substitutions for or additions of amino acids from another FGF molecule. In one embodiment, the C-terminal portion from FGF19 includes a modification that includes a substitution for or addition of amino acid residues from an FGF21 molecule. Exemplary substitutions and additions of such residues are shown in FIGS. 11, 12, and 13.

In one embodiment, the C-terminal portion from FGF19 comprises a modification that includes a substitution of amino acid residues from an FGF21 molecule. In one embodiment, the modification comprises a substitution for or addition of amino acid residues 168 to 209 of SEQ ID NO: 100. In one embodiment, the modification is a substitution of amino acid residues from SEQ ID NO: 100 for corresponding amino acid residues of SEQ ID NO: 1. As shown in FIGS. 5A, 8B, 11, 12, and 13, the corresponding residues of FGF molecules may be identified by sequence analysis and/or structural analysis. In one embodiment, the modification includes a substitution of a contiguous stretch of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, or 42 amino acid residues 168 to 209 of SEQ ID NO: 100 for the corresponding contiguous stretch of amino acid residues of SEQ ID NO: 1. In one embodiment, amino acid residues 169 to 173, 169 to 196, or 169 to 203 of SEQ ID NO: 1 are substituted with the corresponding amino acid residues selected from the sequence comprising amino acid residues 168 to 209 of SEQ ID NO: 100.

In one embodiment, the modification includes a substitution of one or more individual amino acid residues from residues 168 to 209 of SEQ ID NO: 100 for the corresponding amino acid residues of SEQ ID NO: 1. In one embodiment, the C-terminal portion includes substitutions of one or more of amino acid residues 169, 170, 171, 172, 174, 175, 183, 184, 185, 186, 187, 188, 189, 190, 192, 193, 194, 195, 197, 200, 201, 202, 206, 207, 208, 209, 214, 215, or 216 of SEQ ID NO: 1 for the corresponding amino acid residues of SEQ ID NO: 100.

In one embodiment of the present invention, the C-terminal portion from FGF19 includes a modification that includes a deletion of amino acid residues that are absent in the corresponding C-terminal portion from FGF21. As shown in FIGS. 5A, 8B, 11, 12, and 13, FGF19 residues that are absent in the corresponding C-terminal portion of FGF21 may be identified by sequence analysis and/or structural analysis. In one embodiment, the modification comprises a deletion of amino acid residues selected from residues 204 to 216, 197 to 216, 174 to 216, or 169 to 216 of SEQ ID NO: 1. In one embodiment, the modification comprises a deletion corresponding to amino acid residue 204 of SEQ ID NO: 1. In one embodiment, the modification includes a deletion of amino acid residues 178, 179, 180, 181, and/or 182 of SEQ ID NO: 1 individually or in combination.

Chimeric proteins according to the present invention may be isolated proteins or polypeptides. The isolated chimeric proteins of the present invention may be prepared for use in the above described methods of the present invention using standard methods of synthesis known in the art, including solid phase peptide synthesis (Fmoc or Boc strategies) or solution phase peptide synthesis. Alternatively, peptides of the present invention may be prepared using recombinant expression systems.

Accordingly, another aspect of the present invention relates to an isolated nucleic acid molecule encoding a chimeric protein according to the present invention. In one embodiment, the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 211, SEQ ID NO: 212, SEQ ID NO: 213, SEQ ID NO: 214, SEQ ID NO: 215, or SEQ ID NO: 216 (as shown in Table 6). Another aspect of the present invention relates to a nucleic acid construct comprising a nucleic acid molecule encoding a chimeric protein according to the present invention, a 5' DNA promoter sequence, and a 3' terminator sequence. The nucleic acid molecule, the promoter, and the terminator are operatively coupled to permit transcription of the nucleic acid molecule.

TABLE 6

Description of Chimeric Protein Sequence	
Nucleotide sequence of a FGF21/19 chimera composed of	SEQ ID NO: 211 caccccc atccctgact ccagtcctct

TABLE 6-continued

Description of Chimeric Protein Sequence	
residues H29 to V197 of human FGF21 and residues T204 to K216 of human FGF19 (bold)	cctgcaattc gggggccaag tccggcagcg gtacctctac acagatgatg cccagcagac agaagccccac ctggagatca gggaggatgg gacgggtgggg ggcgctgctg acccagagccc cgaaagtctc ctgcagctga aagccttcaa gccccggagtt attcaaataatct tgggagtc gacatccagg ttccctgtgcc agcgcccgaaa tgggggccctg tatggatcgc tccacttga ccctgaggcc tgcaagttcc gggagctgt tcttgaggac ggatacaatg tttaccagtc cgaagccccac ggcctcccgcc tgcacccctg agggaaacaag tccccacacc gggaccctgc accccgagga ccagctcgct tcctgcccact accaggccctg cccccccgcac tcccgagcc accccgaaatc ctggcccccc agcccccccgaa tgtgggctcc tgggaccctc tgacatgg gggactggag gccgtgagga gtcccagtt tgagaagtaa
Nucleotide sequence of a FGF21/19 chimera composed of residues H29 to S190 of human FGF21 and residues M197 to K216 of human FGF19 (bold)	SEQ ID NO: 212 cacccc atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac acagatgatg cccagcagac agaagccccac ctggagatca gggaggatgg gacgggtgggg ggcgctgctg acccagagccc cgaaagtctc ctgcagctga aagccttcaa gccccggagtt attcaaataatct tgggagtc gacatccagg ttccctgtgcc agcgcccgaaa tgggggccctg tatggatcgc tccacttga ccctgaggcc tgcaagttcc gggagctgt tcttgaggac ggatacaatg tttaccagtc cgaagccccac ggcctcccgcc tgcacccctg agggaaacaag tccccacacc gggaccctgc accccgagga ccagctcgct tcctgcccact accaggccctg cccccccgcac tcccgagcc accccgaaatc ctggcccccc agcccccccgaa tgtgggctcc atggaccat ttgggcttgc caccggactg gaggccgtga ggagtcccg ctttgagaag taa
Nucleotide sequence of a FGF21/19 chimera composed of the β-trefoil core domain of human FGF21 (residues H29 to L167) and the C-terminal tail of human FGF19 (residues L169 to K216) (bold)	SEQ ID NO: 213 cacccc atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac acagatgatg cccagcagac agaagccccac ctggagatca gggaggatgg gacgggtgggg ggcgctgctg acccagagccc cgaaagtctc ctgcagctga aagccttcaa gccccggagtt attcaaataatct tgggagtc gacatccagg ttccctgtgcc agcgcccgaaa tgggggccctg tatggatcgc tccacttga ccctgaggcc tgcaagttcc gggagctgt tcttgaggac ggatacaatg tttaccagtc cgaagccccac ggcctcccgcc tgcacccctg agggaaacaag tccccacacc gggaccctgc accccgagga ccagctcgct tcctgcccact actgccccatg gtcccagagg agcctgagga cetcaggggc cacttggaat ctgacatgtt cttctcgccc ctggagaccg acagcatgg ccccatttggg ttgtcacccg gactggggc cgtgaggagt cccagctttt agaagtaa
Nucleotide sequence of a FGF21/19 chimera composed of residues H29 to V197 of human FGF21 harboring Q104M mutation and residues T204 to K216 of human FGF19 (bold)	SEQ ID NO: 214 cacccc atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac acagatgatg cccagcagac agaagccccac ctggagatca gggaggatgg gacgggtgggg ggcgctgctg acccagagccc cgaaagtctc ctgcagctga aagccttcaa gccccggagtt attcaaataatct tgggagtc gacatccagg ttccctgtgcc aatggccaga tgggggccctg tatggatcgc tccacttga ccctgaggcc tgcaagttcc gggagctgt tcttgaggac ggatacaatg tttaccagtc cgaagccccac ggcctcccgcc tgcacccctg agggaaacaag tccccacacc gggaccctgc accccgagga ccagctcgct tcctgcccact accaggccctg cccccccgcac tcccgagcc accccgaaatc ctggcccccc agcccccccgaa

TABLE 6-continued

Description of Chimeric Protein Sequence

Nucleotide sequence of a FGF21/19 chimera composed of residues H29 to S190 of human FGF21 harboring Q104M mutation and residues M197 to K216 of human FGF19 (bold)	SEQ ID NO: 215 tgtgggctcc tcggaccctc tgagcatgg gggactggag gccgtgagga gtcccagctt tgagaagtaa
Nucleotide sequence of a FGF21/19 chimera composed of the β-trefoil core domain of human FGF21 (residues H29 to L167) harboring Q104M mutation and the C-terminal tail of human FGF19 (residues L169 to K216) (bold)	SEQ ID NO: 216 caccgc atccctgact ccagtccctc cctgcaattc gggggccaag tccggcagcg gtaccttac acagatgatg cccacgac agaagccac ctggagatca gggaggatgg gacgggtgggg ggcgtgctg accagagcc cgaaagtctc ctgcagctga aagcattgaa gccccggattt attcaaatct tgggagtc gacatccagg ttccctgtgcc aatggccaga tggggccctg tatggatcgc tccacttta ccctgaggcc tgcagttcc gggagctgt tcttgaggac ggatacaatg tttaccatc cgaagccac ggcctcccgcc tgcacccgt agggaaacaag tccccacacc gggaccctgc accccggagga ccagctcgat tccctgca accaggccctg cccccccgcac tcccgagcc acccggaaatc ctggcccccc agccccccga tgtgggctcc atggaccat ttggcccttg cacccggactg gaggccgtga ggagtcctcag ctttgagaag taa

Also encompassed are vectors or expression vectors comprising such nucleic acid molecules and host cells comprising such nucleic acid molecules. Nucleic acid molecules according to the present invention can be expressed in a host cell, and the encoded polynucleotides isolated, according to techniques that are known in the art.

Generally, the use of recombinant expression systems involves inserting the nucleic acid molecule encoding the amino acid sequence of the desired peptide into an expression system to which the molecule is heterologous (i.e., not normally present). One or more desired nucleic acid molecules encoding a peptide of the invention may be inserted into the vector. When multiple nucleic acid molecules are inserted, the multiple nucleic acid molecules may encode the same or different peptides. The heterologous nucleic acid molecule is inserted into the expression system or vector in proper sense (5'→3') orientation relative to the promoter and any other 5' regulatory molecules, and correct reading frame.

The preparation of the nucleic acid constructs can be carried out using standard cloning procedures well known in the art as described by Joseph Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL (Cold Springs Harbor 1989). U.S. Pat. No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference in its entirety, describes the production of expression systems in the form of recombinant

plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in a suitable host cell.

A variety of genetic signals and processing events that control many levels of gene expression (e.g., DNA transcription and messenger RNA ("mRNA") translation) can be incorporated into the nucleic acid construct to maximize protein production. For the purposes of expressing a cloned nucleic acid sequence encoding a desired protein, it is advantageous to use strong promoters to obtain a high level of transcription. Depending upon the host system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, lac promoter, trp promoter, recA promoter, ribosomal RNA promoter, the P_R and P_L promoters of coliphage lambda and others, including but not limited, to lacUV5, ompF, bla, lpp, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid trp-lacUV5 (lac) promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene. Common promoters suitable for directing expression in mammalian cells include, without limitation,

103

SV40, MMTV, metallothionein-1, adenovirus E1a, CMV, immediate early, immunoglobulin heavy chain promoter and enhancer, and RSV-LTR.

There are other specific initiation signals required for efficient gene transcription and translation in prokaryotic cells that can be included in the nucleic acid construct to maximize protein production. Depending on the vector system and host utilized, any number of suitable transcription and/or translation elements, including constitutive, inducible, and repressible promoters, as well as minimal 5' promoter elements, enhancers or leader sequences may be used. For a review on maximizing gene expression see Roberts and Lauer, "Maximizing Gene Expression On a Plasmid Using Recombination *In Vitro*," *Methods in Enzymology* 68:473-82 (1979), which is hereby incorporated by reference in its entirety.

A nucleic acid molecule encoding an isolated protein of the present invention, a promoter molecule of choice, including, without limitation, enhancers, and leader sequences; a suitable 3' regulatory region to allow transcription in the host, and any additional desired components, such as reporter or marker genes, are cloned into the vector of choice using standard cloning procedures in the art, such as described in Joseph Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL* (Cold Springs Harbor 1989); Frederick M. Ausubel, *SHORT PROTOCOLS IN MOLECULAR BIOLOGY* (Wiley 1999); and U.S. Pat. No. 4,237,224 to Cohen and Boyer, which are hereby incorporated by reference in their entirety.

Once the nucleic acid molecule encoding the protein has been cloned into an expression vector, it is ready to be incorporated into a host. Recombinant molecules can be introduced into cells, without limitation, via transfection (if the host is a eukaryote), transduction, conjugation, mobilization, or electroporation, lipofection, protoplast fusion, mobilization, or particle bombardment, using standard cloning procedures known in the art, as described by JOSEPH SAMBROOK et al., *MOLECULAR CLONING: A LABORATORY MANUAL* (Cold Springs Harbor 1989), which is hereby incorporated by reference in its entirety.

A variety of suitable host-vector systems may be utilized to express the recombinant protein or polypeptide. Primarily, the vector system must be compatible with the host used. Host-vector systems include, without limitation, the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria.

Purified proteins may be obtained by several methods readily known in the art, including ion exchange chromatography, hydrophobic interaction chromatography, affinity chromatography, gel filtration, and reverse phase chromatography. The protein is preferably produced in purified form (preferably at least about 80% or 85% pure, more preferably at least about 90% or 95% pure) by conventional techniques. Depending on whether the recombinant host cell is made to secrete the protein into growth medium (see U.S. Pat. No. 6,596,509 to Bauer et al., which is hereby incorporated by reference in its entirety), the protein can be isolated and purified by centrifugation (to separate cellular components from supernatant containing the secreted protein) followed by sequential ammonium sulfate precipitation of the supernatant. The fraction containing the protein is subjected to gel filtration in an appropriately sized dextran or

104

polyacrylamide column to separate the protein of interest from other proteins. If necessary, the protein fraction may be further purified by HPLC.

Another aspect of the present invention relates to a pharmaceutical composition that includes a chimeric protein according to the present invention and a pharmaceutically acceptable carrier.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are non-toxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; 10 antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICSTM.

The term "pharmaceutically acceptable" means it is, within the scope of sound medical judgment, suitable for use in contact with the cells of humans and lower animals without undue toxicity, irritation, allergic response, and the like, and is commensurate with a reasonable benefit/risk ratio.

In one embodiment, the chimeric protein of the present invention or a pharmaceutical composition thereof is administered in a therapeutically effective amount in combination with a therapeutically effective amount of a second agent. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof is administered in conjunction with the second agent, i.e., the respective periods of administration are part of a single administrative regimen. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof and the second agent are administered concurrently, i.e., the respective periods of administration overlap each other. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof and the second agent are administered non-concurrently, i.e., the respective periods of administration do not overlap each other. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof and the second agent are administered sequentially, i.e., the chimeric protein of the present invention or pharmaceutical composition thereof is administered prior to and/or after the administration of the second agent. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof and the second agent are administered simultaneously as separate compositions. In one embodiment, the chimeric protein of the present invention or pharmaceutical composition thereof and the second agent are administered simultaneously as part of the same compositions.

In one embodiment, the second agent is an anti-inflammatory agent, an antihypertensive agent, an anti-diabetic agent, and/or cholesterol-lowering drug such as a drug of the "statin" class. In one embodiment, the second agent is insulin. In one embodiment, the insulin is rapid acting, short acting, regular acting, intermediate acting, or long acting insulin. In one embodiment, the insulin is and/or comprises Humalog, Lispro, Novolog, Apidra, Humulin, Aspart, regu-

105

lar insulin, NPH, Lente, Ultralente, Lantus, Glargine, Levemir, or Detemir. In one embodiment, the second agent is a statin. In one embodiment, the statin is and/or comprises Atorvastatin (e.g., Lipitor or Torvast), Cerivastatin (e.g., Lipobay or Baycol), Fluvastatin (e.g., Lescol or Lescol), Lovastatin (e.g., Mevacor, Altocor, or Altoprev) Mevacor, Pitavastatin (e.g., Livalo or Pitava), Pravastatin (e.g., Pravachol, Selektine, or Lipostat) Rosuvastatin (e.g., Crestor), Simvastatin (e.g., Zocor or Lipex), Vytorin, Advicor, Besylate Caduet or Simcor.

In one embodiment of the present invention, the pharmaceutical composition according to the present invention is administered with an anti-inflammatory agent, an antifibrotic agent, an antihypertensive agent, an antidiabetic agent, a triglyceride-lowering agent, and/or a cholesterol-lowering agent.

Another aspect of the present invention relates to a method of treating a subject suffering from diabetes, obesity, or metabolic syndrome. This method includes selecting a subject suffering from diabetes, obesity, or metabolic syndrome and administering to this selected subject a therapeutically effective amount of a chimeric protein according to the present invention.

In one embodiment, the selected subject is a mammal. In one particular embodiment, the selected subject is a human. In another embodiment, the selected subject is a rodent.

In one embodiment the selected subject has diabetes. As used herein, diabetes includes, but is not limited to, type I diabetes, type II diabetes, gestational diabetes, and drug-induced diabetes. In one embodiment, the subject has obesity. In one embodiment, the subject has metabolic syndrome.

The pharmaceutical compositions comprising a chimeric protein of the present invention provided herein can be used to treat a number of conditions. Preferably, the condition is one which the therapeutic outcome includes a decrease in blood glucose, a decrease in blood fructosamine, an increase in energy expenditure, an increase in fat utilization, a decrease in body weight, a decrease in body fat, a decrease in triglycerides, a decrease in free fatty acids, an increase in fat excretion, an improvement, or even a preservation, of pancreatic β -cell function and mass, a decrease in total blood cholesterol, a decrease in blood low-density lipoprotein cholesterol, an increase in blood high-density lipoprotein cholesterol, an increase in blood adiponectin, an increase in insulin sensitivity, an increase in leptin sensitivity, a decrease in blood insulin, a decrease in blood leptin, a decrease in blood glucagon, an increase in glucose uptake by adipocytes, a decrease in fat accumulation in hepatocytes, and/or an increase in fat oxidation in hepatocytes. Each of these parameters can be measured by standard methods, for example, by measuring oxygen consumption to determine metabolic rate, using scales to determine weight, and measuring lean body mass composition or mass to determine fat. Moreover, the presence and amount of triglycerides, free fatty acids, glucose and leptin can be determined by standard methods (e.g., blood test).

Additional conditions that are treatable in accordance with the present invention include one or more of type 1 diabetes, type 2 diabetes, gestational diabetes, drug-induced diabetes, high blood glucose, metabolic syndrome, lipodystrophy syndrome, dyslipidemia, insulin resistance, leptin resistance, atherosclerosis, vascular disease, inflammatory disease, fibrotic disease, hypercholesterolemia, hypertriglyceridemia, non-alcoholic fatty liver disease, overweight, and obesity.

106

The pharmaceutical composition according to the present invention can be administered orally, parenterally, subcutaneously, intravenously, intramuscularly, intraperitoneally, by intranasal instillation, by implantation, by intracavitary or intravesical instillation, intraocularly, intraarterially, intralesionally, transdermally, or by application to mucous membranes. The most suitable route may depend on the condition and disorder of the recipient. The formulations may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of pharmacy.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Those skilled in the art can readily optimize pharmaceutically effective dosages and administration regimens for therapeutic compositions comprising the chimeric protein according to the present invention, as determined by good medical practice and the clinical condition of the individual patient.

When *in vivo* administration of a chimeric protein of the present invention or is employed, normal dosage amounts may vary from, for example, about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day. In one embodiment, the dosage may be from about 1 mg/kg/day to 10 mg/kg/day, depending upon the route of administration. In one embodiment, the chimeric protein according to the present invention is administered at a dose of about 0.1 to 10 mg/kg once or twice daily. In one embodiment, the chimeric protein according to the present invention is administered at a dose of about 1 to 9, 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, or 1 to 2 mg/kg. In one embodiment, the dosage is the same as that of a native FGF21 therapeutic. In one embodiment, the dosage is less than that of a native FGF21 therapeutic, but having the same effect as a higher dosage of a native FGF21 therapeutic. Guidance as to particular dosages and methods of delivery of proteins is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212, which are hereby incorporated by reference in their entirety. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a chimeric protein of the present invention is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the chimeric protein of the present invention, microencapsulation is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., "Preparation and Characterization of Poly(D,L-lactide-co-glycolide) Microspheres for Controlled Release of Human Growth Hormone," *Nat. Med.* 2:795-799 (1996); Yasuda, "Sustained Release Formulation of Interferon," *Biomed. Ther.* 27:1221-1223 (1993); Hora et al., "Controlled Release of Interleukin-2 from Biodegradable Microspheres," *Nat. Biotechnol.* 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in VACCINE DESIGN: THE SUB-UNIT AND ADJUVANT APPROACH 439-462 (Powell and Newman, eds. 1995); WO 97/03692; WO 96/40072; WO 96/07399; and U.S. Pat. No. 5,654,010, which are hereby incorporated by reference in their entirety. The sustained-release formu-

107

lations of these proteins were developed using poly-lactic-glycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: BIODEGRADABLE POLYMERS AS DRUG DELIVERY SYSTEMS 1-41 (M. Chasin and R. Langer eds. 1990), which is hereby incorporated by reference in its entirety.

The compositions according to the present invention may be administered as frequently as necessary in order to obtain the desired therapeutic effect. Some patients may respond rapidly to a higher or lower dose and may find much weaker maintenance doses adequate. For other patients, it may be necessary to have long-term treatments at the rate of 1 to 4 doses per day, in accordance with the physiological requirements of each particular patient. For other patients, it will be necessary to prescribe not more than one or two doses per day.

Another aspect of the present invention relates to a method of treating a subject in need of increased FGF21- β Klotho-FGF receptor ("FGFR") complex formation. This method includes selecting a subject in need of increased FGF21- β Klotho-FGFR complex formation and administering to the selected subject a chimeric FGF21 protein, where the chimeric FGF21 protein includes an FGF21 core domain and a C-terminal portion of FGF19, thereby treating a subject in need of increased FGF21- β Klotho-FGFR complex formation.

Suitable chimeric FGF21 proteins include chimeric proteins according to the present invention, as described above and throughout the present application.

108

5 FGF21 depends on the co-receptor β Klotho to activate its cognate FGFR (FGFR1c) in its target tissues including white adipose tissue (Ogawa et al., "BetaKlotho is Required for Metabolic Activity of Fibroblast Growth Factor 21," *Proc. Natl. Acad. Sci. USA* 104(18):7432-7437 (2007), which is hereby incorporated by reference in its entirety). In the course of deciphering the molecular details of how FGF21 forms a signaling complex on the cell surface with FGFR1c and β Klotho, two discoveries were made that provided the basis for the rational design of an FGF21 agonist. It was found that β Klotho promotes binding of FGF21 to its cognate FGFR by engaging ligand and receptor simultaneously through two distinct binding sites. β Klotho plays the same role in promoting binding of FGF19, an endocrine regulator of bile acid homeostasis, to its cognate FGFR. The binding site for β Klotho was mapped on FGF21 and FGF19 to the C-terminal region of each ligand that follows the β -trefoil core domain. In the course of these studies, it was found that the C-terminal tail peptides of FGF21 and FGF19 share a common binding site on β Klotho, and that the 10 C-terminal tail of FGF 19 binds tighter than the C-terminal tail of FGF21 to this site. Based on these findings, chimeric FGF21 proteins were made in which C-terminal sequences in FGF21 were replaced with the corresponding sequences of FGF 19 in order to confer greater binding affinity to β Klotho, and enhance agonistic properties.

15 In one embodiment according to the present invention, β Klotho is mammalian β Klotho. In one embodiment, β Klotho is human or mouse β Klotho. In one particular embodiment of the present invention, β Klotho is human or mouse β Klotho comprising the amino acid sequence of SEQ ID NO: 217 (i.e., GenBank Accession No. NP_783864, which is hereby incorporated by reference in its entirety) or SEQ ID NO: 218 (i.e., GenBank Accession No. NP_112457, which is hereby incorporated by reference in its entirety), respectively, as follows:

```

SEQ ID NO: 217:
1 MKPGCAAGSP GNEWIFFSTD EITTRYRNTM SNGGLQRSVI LSALILLRAV TGFSGDGRAI
61 WSKNPNFTPV NESQLFLYDT FPKNFFWIGIG TGALQVEGSW KKDGKGPSIW DHFIHTHLKN
121 VSSTNGSSD YIFLEKDLSA LDFIGVSFYQ FSISWPRLFP DGIVTVANAK GLQYYSTLLD
181 ALVLRNIEPI VTLYHWDLPL ALQEKYGGWK NDTIIDIFND YATYCFQMFG DRVKYWITIH
241 NPYLVVAHGY GTGMHAPGEK GNLAAYVTVG HNLKAHSKV WHNYNTHFRP HQKGWLSTIL
301 GSHWIEPNRS ENTMDIFKCQ QSMVSVLGWF ANPIHGDDY PEGMRKKLFS VLPIFSEAEK
361 HEMRGTDAFF AFSFGPNNFK PLNTMAKMGQ NVSLNLREAL NWIKLEYNNP RILIAENGWF
421 TDSRVKTEDT TAIYMMKNFL SQVLQAIRLD EIRVFGYTAW SLLEGFEWQD AYTIRRLGLFY
481 VDFNSKQKER KPKSSAHYYK QIIRENGFSL KESTPDVQGQ FPCDFSWGVT ESVLKPESEA
541 SSPQFSDPHL YVWNATGNRL LHRVEGVRLK TRPAQCTDFV NIKKOLEMLA RMKVTHYRFA
601 LDWASVLPNG NLSAVNRQAL RYYRCVVSEG LKLGISAMVT LYYPTHAHLG LPEPLLHADG
661 WLNPSTAEEAF QAYAGLCFQE LGDLVKLWIT INEPNRLSDI YNRSGNDTYG AAHNLLVAHA
721 LAWRLYDRQF RPSQRGAWSL SLHADWAEPN NPYADSHWRA AERFLQFEIA WFAEPLFKTG
781 DYPAAMREYI ASKHRRGLSS SALPRLTEAE RRLLKGTVDF CALNHFTTRF VMHEQLAGSR
841 YDSDRDIQFL QDITRLSSPT RLAVIPWGVR KLLRWVRRNY GDMDIYITAS GIDDQALED
901 RLRKYYLGKY LQEVLKAYLI DKVRIKGYYA FKLAEEKSKP RFGFFTSDFK AKSSIQFYNK
961 VISSRGFPFE NSSSRCSTQ ENTECTVCLF LVQKKPLIFL GCCFFSTLVL LLSIAIFQRQ
1021 KRRKFWKAKN LQHPLKKGK RVVS

```

- continued

SEQ ID NO: 218:

```

1 MKTGCAAGSP GNEWIFFSSD ERNTRSRTKM SNRALQRSAV LSAFVLLRAV TGFSGDGKAI
61 WDKKQVSPV NPSQLFLYDT FPKNFSWVG VG TGAFQVEGSW KTDGRGPSIW DRYVYSHLRG
121 VNGTDRSTD YIFLEKDLLA LDFLGVSFYQ FSISWPRLFP NGTVAAVNAQ GLRYYRALLD
181 SLVLRNIEPI VTLYHWDLPL TLQEEYGGWK NATMIDLFDN YATYCFQTFG DRVKYWITIH
241 NPYLVAWHGF GTGMHAPGEK GNLTAVYTVG HNLIKAHSKV WHNYDKNFRP HQKGWLTSITL
301 GSHWIEPNRT DNMEDVINCO HSMSSVLGWF ANPIHGDDY PEFMKTGAMI PEFSEAEKEE
361 VRGTADFFAF SFGPNNFRPS NTVVKMGQNV SLNLRQVLNW IKLEYDDPQI LISENGWFTD
421 SYIKTEDTTA IYMMKNFLNQ VLQAIKFDEI RVFGYTAWTL LDGF EWQDAY TTRRGLFYVD
481 FNSEQKERKP KSSAHYYKQI IQDNGFPLKE STPDPMKGRFP CDFSWGVTES VLKPEFTVSS
541 PQFTDPHLYV WNVTGNRLLY RVEGVRLKTR PSQCTDYVSI KKRVEMLAKM KVTHYQFALD
601 WTSILPTGNL SKVNRQVLRY YRCVVSEGLK LGVFPMTLY HPTHSHLGLP LPLLSSGGWL
661 NMNTAKAFQD YAELCFRELG DLVKLWITIN EPNRLSDMYN RTSNDTYRAA HNLMIAHAQV
721 WHLYDRQYRP VQHGAVSLSL HCDWAEPANP FVDSHWKAAE RFLQFEIAWF ADPLFKTGDY
781 PSVMKEYIAS KNQRGLSSSV LPRPTAKESR LVKGTVDFYA LNHFTRFVI HKQLNTNRSV
841 ADRDVQFLQD ITRLSSPSRL AVTPWGVRKL LAWIRRNYRD RDIYITANGI DDLALEDDQI
901 RKYYLEKYVQ EALKAYLIDK VKIKGYYAFK LTEEKSKPRF GFFTSDFRAK SSVQFYSKLI
961 SSSGLPAENR SPACGQPAED TDCTICSLV EKKPLIUFFGC CFISTLAVLL SITVPHHQKR
1021 RKFQKARNLQ NIPLKKGHSR VFS

```

In one particular embodiment of the present invention, βKlotho is human or mouse βKlotho encoded by a nucleotide sequence comprising the nucleotide sequences of SEQ ID NO: 219 (GenBank Accession No. NM_175737, which

is hereby incorporated by reference in its entirety) and SEQ ID NO: 220 (GenBank Accession No. NM_031180, which is hereby incorporated by reference in its entirety), as follows:

SEQ ID NO: 219 (Human βKlotho gene coding sequence):

```

98      atg aagccaggct gtgcggcagg atctccagg aatgaatggaa ttttcttcag
151     cactgatgaa ataaccacac gctataggaa tacaatgtcc aacggggat tgcaaagtc
211     tgtcatcctg tcagcactta ttctgtctacg agctgttact ggattctctg gagatggaa
271     agctatatgg tctaaaaatc ctaattttac tccggtaaat gaaagttagc tgtttctcta
331     tgacacttc cctaaaaact ttttctgggg tattggact ggagcattgc aagtggaaagg
391     gagttggaaag aaggatggaa aaggaccttc tatatggat catttcattcc acacacacct
451     taaaaatgtc agcagcacga atgggtccag tgacagttat attttctgg aaaaagactt
511     atcagccctg gatttatag gagtttcttt ttatcaattt tcaatttcct ggccaaggct
571     ttccccatg ggaatagtaa cagttggca cgcaaaagggt ctgcgtact acagtactct
631     tctggacgct ctagtgctta gaaacattga acctatacg actttatacc actgggattt
691     gcctttggca ctacaagaaa aatatgggg gtggaaaaat gataccataa tagatatctt
751     caatgactat gocacatact gttccagat gttggggac cgtgtcaaattt attggattac
811     aattcacaac ccataatctag tggcttggca tgggtatggg acaggtatgc atgcccctgg
871     agagaaggaa aatttacgacat ctgtctacac tggggacac aacttgcatac aggctcactc
931     gaaagttgg cataactaca acacacattt ccgcccacat cagaagggtt ggttatcgat
991     cacgttggaa tctcatttgc tcgagccaaa ccggctggaa aacacgatgg atatattcaa
1051    atgtcaacaa tccatggttt ctgtgttgc aaccctatcc atggggatgg

```

111**112**

-Continued

1111 cgactatcca gaggggatga gaaagaagtt gttctccgtt ctacccattt tctctgaagc
 1171 agagaagcat gagatgagag gcacagctga tttcttgcc ttttcttttgc gaccaacaa
 1231 cttcaagccc ctaaacacca tggctaaaat gggacaaaat gtttcactta atttaagaga
 1291 agcgctgaac tggattaaac tggataacaa caaccctcg aatcttgcattt ctgagaatgg
 1351 ctggttcaca gacagtctgt tgaaaacaga agacaccacg gccatctaca tgatgaagaa
 1411 ttccctcagc cagggtcttc aagcaataag gtttagatgaa atacgagtgt ttggttatac
 1471 tgcctggatc ctccctggatc gctttgaatg gcaggatgct tacaccatcc gccgaggatt
 1531 attttatgtg gattttaca gtaaacagaa agagcggaaa cctaagtctt cagcacacta
 1591 ctacaaacag atcatacgg aaaaatggttt ttctttaaaa gagtccacgc cagatgtgca
 1651 gggccagttt ccctgtact tctccctgggg tgtcaactgaa tctgttctt aagcccgagtc
 1711 tgtggctcg tccccacagt tcacgcgtcc tcacatgtac gtgtggaaacg ccactggcaa
 1771 cagactgttg caccgagtgg aagggggttag gctgaaaaca cgacccgctc aatgcacaga
 1831 ttttgtaaac atcaaaaaac aacttggatg gttggcaaga atgaaaagtca cccactaccg
 1891 gtttgctctg gattgggcct cggtccttcc cactggcaac ctgtccgcgg tgaaccgaca
 1951 ggccctgagg tactacaggc gctgtggctag tgaggggctg aagcttggca tctccgcgt
 2011 ggtcacccctg tattatccga cccacgcccc cctaggcctc cccgagccctc tggatgc
 2071 cgacgggtgg ctgaacccat cgacggccga ggccttccag gcctacgctg ggctgtgctt
 2131 ccaggagctg ggggacctgg tgaagctctg gatcaccatc aacgagccctaa accggctaaag
 2191 tgacatctac aaccgcctcg gcaacgacac ctacggggcg ggcacaaccc tgctgggtggc
 2251 ccacgcctg gcctggcgcc tctacgaccg gcaagttcagg ccctcacagc gccccccgt
 2311 gtcgctgtcg ctgcacgcgg actgggcggaa acccgccaaac ccctatgctg actcgactg
 2371 gaggggggcc gagcgttcc tgcagttcga gatgcctgg ttcggcggc cgctttca
 2431 gacgggggac taccggcgcc ccatgaggaa atacattgcc tccaaaggcacc gacgggggct
 2491 ttccagctcg gcctgcgcg gctcaccga ggccgaaagg aggtgtgtca agggcacgg
 2551 cgacttcgtc ggcgtcaacc acattcaccac taggttcgtg atgcacgagc agctggccgg
 2611 cagccgtac gactcggaca gggacatcca gtttctgcag gacatcaccc gcctgagctc
 2671 cccacgcgc ctggctgtga ttccctgggg ggtgcgcga ctgcgtcggt gggccggag
 2731 gaactacggc gacatggaca tttacatcac cgccagtgcc atcgacgacc aggctctgga
 2791 ggtatgacccgg ctccggaaatg actacatgggaaatg tacatggatc caggaggtgc tgaaaagcata
 2851 cctgattgt aaagtca gaa tcaaaggcta ttatgcattt aaactggctg aagagaaatc
 2911 taaacccaga ttggattct tcacatctga ttttaaagttt aaatcctcaa tacaattttta
 2971 caacaaatgt atcagcagca ggggcttccc ttttggaaac agtagttcta gatgcgtca
 3031 gacccaaagaa aatacagatg gcactgtctg cttatccctt gtgcagaaga aaccactgt
 3091 attccctgggt ttttttttttct tctccacccct gtttctactc ttatcaatttgc ccattttca
 3151 aaggcagaag agaagaaaatg tttggaaagc aaaaaacttca caacacatac cattaaagaa
 3211 aggcaagaga gttgttagct aa

SEQ ID NO: 220 (House mouse β Klotho gene coding sequence):

2 atgaagaca ggctgtgcag cagggtctcc gggaaatgaa tggattttct tcagctctga
 61 tggatggaaaac acacgcgtctca gggaaacaaat gtccaaacagg gcactgcaaa gatctggccgt
 121 gctgtctgcg tttgttctgc tgcgagctgt taccggcttc tccggagacg gggaaagcaat
 181 atgggataaa aaacagtcacg tgagtccggtaaaacccaaatg cagctgttcc tctatgacac

-Continued

241 tttccctaaa aactttctt gggcggttg gaccggaca tttcaagtgg aaggaggttg
 301 gaagacagat ggaaggaggac cctcgatctg ggatcggta cgtctactcac acctgagagg
 361 tgtcaacggc acagacagat ccactgacag ttacatcttt ctgaaaaaag acttgttgc
 421 tctggatttt ttaggagttt cttttatca gttctcaatc tcctggccac gggtgttcc
 481 caatgaaaca gtagcagcag tgaatgcga aggtctccgg tactaccgtg cacttctgga
 541 ctcgctggta ctttagaata tcgagccccat tgtagcttg taccattggg atttgctct
 601 gacgctccag gaagaatatg ggggctggaa aaatgcaact atgatagatc tcttcaacga
 661 ctatgccaca tactgcttcc agaccccttgg agaccgtgtc aaatattggta ttacaattca
 721 caacccttac ctttgtgctt ggcatgggtt tggcacaggt atgcacatgcac caggagagaa
 781 gggaaattta acagctgtct acactgtggg acacaacctg atcaaggcac attcgaaagt
 841 gtggcataac tacgacaaaaa acttccggccc tcatcagaag ggttggctt ccatcacctt
 901 ggggtcccat tggatagagc caaacagaac agacaacatg gaggacgtga tcaactgcca
 961 gcactccatg tcctctgtgc ttggatgggtt cgccaaccccc atccacgggg acggcacta
 1021 ccctgagttc atgaagacgg ggcgcattat ccccgagttc tctgaggcag agaaggagga
 1081 ggtgaggggc acggctgatt tcttgcctt ttccctcggg cccaaacaact tcaggccctc
 1141 aaacaccgtg gtgaaaatgg gacaaaatgt atcactcaac ttaaggcagg tgctgaactg
 1201 gattaaactg gaatacgtg accctcaaattt ctgttgcctg gagaacggct ggttcacaga
 1261 tagctatata aagacagagg acaccacggc catctacatg atgaagaatt tcctaaacca
 1321 ggttcttcaa gcaataaaaat ttgatgaaat ccgcgtgttt ggttatacgg cctggactct
 1381 cctggatggc tttgagtgcc aggatgccta tacgacccga cgagggtgt tttatgtgga
 1441 cttaacagt gacgacaaaaa agaggaaacc caagtccatcg gtcattact acaacgagat
 1501 catacaagac aacggcttcc ctgttgcctt gtcacgcaca gacatgaagg gtcgggttccc
 1561 ctgtgatttc ctgttggggag tcactgagtc tggttcttaa cccgagttt cgggttcctc
 1621 ccccgagttt accgatecctc acctgtatgt gtggaatgtc actggcaaca gattgtctca
 1681 ccgagtgaa ggggtaaaggc tgaaaacaag accatcccag tgacacagatt atgtgagcat
 1741 caaaaaacga gttgaaaatgt tggaaaaat gaaagtccacc cactaccgtg ttgtctgga
 1801 ctggacccatc atcccttccca ctggcaatct gtccaaaggta aacagacaag tggtaaggta
 1861 ctataggtgt gtggtgagcg aaggactgaa gctggcgctc ttccctatgg tgacgttgta
 1921 ccacccaacc cactccatc tcggcctccccc cctggccactt ctgagcgtg ggggggtggct
 1981 aaacatgaac acagccaaagg cttccagga ctacgctgag ctgtgtttcc gggagttggg
 2041 ggacttggtg aagctctggta tcaccatcaa tgagccatac aggctgagtg acatgtacaa
 2101 ccgcacagat aatgacacactt accgtgcagc ccacaacccgt atgatgcggcc atgcccaggt
 2161 ctggcacctc tatgataggc agtataaggcc ggtccagcat ggggtgtgt cgctgtccctt
 2221 acattgcgac tggcagaac ctggcaaccc ctgttgcgtt tcacactgga aggcagccga
 2281 ggcgttcctc cagtttgaga tcgcctgggtt tgcaatccg ctcttcaaga ctggcgacta
 2341 tccatcggtt atgaaggaat acatgcctc caagaaccag cgagggtgt ctatgtcgt
 2401 cctggccgcgc ttccacccgca aggagagcag gctggtaag ggttaccgtcg acttctacgc
 2461 actgaaccac ttcaactacga gggtcgat acacaacccgatc ctgaacacca accgtcgat
 2521 tgcagacagg gacgtccatg tccatgcgtt catcaccgc ctaagctgc ccageccgcct
 2581 ggctgttaaca ccctggggag tgccaaagct ctttgcgtgg atccggagga actacagaga
 2641 caggatatac tacatcacag ccaatggcat cgatgacccgtg gtcgttaggg atgatcgat

-continued

2701 ccgaaagtac tacttgaga agtatgtcca ggaggctctg aaagcatatc tcattgacaa
 2761 ggtcaaaaatc aaaggctact atgcattcaa actgaactgaa gagaaatcta agcctagatt
 2821 tggattttc acctctgact tcagagctaa gtcctctgtc cagtttaca gcaagctgat
 2881 cagcagcagt ggcctccccg ctgagaacag aagtccctcg tgggtcagc ctgcggaga
 2941 cacagactgc accatttgtc catttctgtc ggagaagaaa ccactcatct tcttcgggttgc
 3001 ctgcttcatc tccactctgg ctgtactgtc atccatcacc gttttcatc ataaaaagag
 3061 aagaaaattc cagaaagcaa ggaacttaca aaatatacca ttgaagaaag gccacagcag
 3121 agtttcaagc taa

15

In one embodiment of the present invention, the FGF receptor is FGFR1c receptor. In one particular embodiment, the FGFR1c receptor is the human FGFR1c receptor com-

prising the amino acid sequence of SEQ ID NO: 221 (GenBank Accession No. NP_075598, which is hereby incorporated by reference in its entirety), as follows:

```

1 MWSKCLLFW AVLVTATLCT ARPSPTLPEQ AQPWGAPVEV ESFLVHPGDL LQLRCRLRDD
  61 VQSINWLRDG VQLAESNRTR ITGEEEVQD SVPADSGLYA CVTSSPSGSD TTYFSVNVD
  121 ALPSSEDDDD DDDSSSEEKE TDNTKPNRMP VAPYWTTSPEK MEKKLHAVPA AKTVKFKCPS
  181 SGTPNPLRW LKNGKEFPD HRIGGYKVRY ATWSIIMDSV VPSDKGNYTC IVENEYGSIN
  241 HTYQLDVVER SPHRPILQAG LPANKTVALG SNVEFMCKVY SDPQPHIQWL KHIEVNGSKI
  301 GPDNLPLYVQI LKTAGVNNTD KEMEVLHLRN VSFEDAGEYT CLAGNSIGLS HHSAWLTVLE
  361 ALEERPAVMT SPLYLEIIIIY CTGAFLISCM VGSVIVYKMK SGTKKSDFH QMAVHKLAKS
  421 IPLRRQVTVS ADSSASMNSG VLLVRPSRLS SSGTPMLAGV SEYELPEDPR WELPRDRLV
  481 GKPLGEGCFG QVVLAEAIGL DKDKPNRVTK VAVKMLKSDA TEKDLSDLIS EMEMMKMIGK
  541 HKNIINLLGA CTQDGPLYVI VEYASKGNLR EYLQARRPPG LEYCYNPSHN PEEQLSSKDL
  601 VSCAYQVARG MEYLASKKCI HRDLAARNVL VTEDNVMKIA DFGLARDIH IDYYKTTNG
  661 RLPVKWMAPE ALFDRIYTHQ SDVWSFGVLL WEIFTLGGSP YPGVPVEELF KLLKEGHMRD
  721 KPSNCTNELY MMRDCWHAV PSQRPTFKQL VEDLDRIVAL TSNQEYLDLS MPLDQYSPSF
  781 PDTRSSTCSS GEDSVFSHEP LPPEPCLPRH PAQLANGGLK RR

```

In one particular embodiment of the present invention, the
 45 FGFR1c receptor is the human FGFR1c receptor encoded by a nucleotide sequence comprising the nucleotide sequence of SEQ ID NO: 222 (GenBank Accession No. NM_023110, which is hereby incorporated by reference in its entirety), as follows:

SEQ ID NO: 222 (Human FGFR1c gene coding sequence) :

```

  943 atgtggag ctggaaagtgc ctccctttctt gggctgtgtc ggtcacagcc acactctgca
  1001 ccgcttaggcc gtcggcggacc ttgcctgaac aagcccgacc ctggggagcc cctgtggaaag
  1061 tggagtcctt cctgggtccac cccgggtgacc tgctgcagct tgcgtgtcgg ctgcggggacg
  1121 atgtgcagag catcaactgg ctgcgggacg gggtgccagct ggccggaaagc aaccgcaccc
  1181 gcatcacagg ggaggagggtg gaggtgcagg actccgtgcc cgccagactcc ggcctctatg
  1241 cttgcgtaac cagcagcccc tcggggcagtgc acaccaccta cttctccgtc aatgtttcag
  1301 atgtctccc ctcctcgag gatgtatgtc atgtatgtatgc ctcctttca gaggagaaag
  1361 aaacagataa caccaaacca aaccgttatgc ccgtagctcc atatggaca tccccagaaa
  1421 agatggaaaa gaaattgcat gcagtgcgg ctgccaagac agtgaagtcc aatgcctt

```

-Continued

1481 ccagtggac cccaaacccc acatgcgt ggttaaaaaa tggcaaagaa ttcaaacctg
 1541 accacagaat tggaggctac aaggccgtt atgccacctg gagcatcata atggactctg
 1601 tggtcctc tgacaaggc aactacacct gcattgtga gaatgactac ggcagcatca
 1661 accacacata ccagctggat gtcgtggc ggtccctca ccggcccatc ctgcaagcag
 1721 gggtccccgc caacaaaaca gtggccctgg gtagcaacgt ggagttcatg tgtaaggtgt
 1781 acagtgaccc gcagccgcac atccagtggc taaagcacat cgaggtgaat gggagcaaga
 1841 ttggccaga caacctgc tatgtccaga tcttgaagac tgctggagtt aataccaccc
 1901 acaaagat ggaggtgctt cacttaagaa atgtctcctt tgaggacgca ggggagtata
 1961 cgtgcttggc gggtaactct atcgactct cccatcaactc tgcatggttg accgttctgg
 2021 aagccctgga agagaggccc gcagtgatga cctcgccctt gtacctggag atcatcatct
 2081 attgcacagg ggccttcctc atctcctgc tggggggc ggtcatcgta tacaagatga
 2141 agagtggatc caagaagat gacttccaca gccagatggc tgtgcacaag ctggccaaga
 2201 gcatccctct ggcgacacag gtaacagtgt ctgctgactc cagtgcaccc atgaactctg
 2261 gggttcttctt ggttcggcca tcacggctct cctccagtg gactccatg ctacgggg
 2301 tctctgagta tgagctccc gaagaccctc gctggagct gcctcggac agactggct
 2361 taggcaaacc cctggggagag ggctgcttg ggcaggtgtt gttggcagag gctatcggc
 2421 tggacaagga caaaccacac cgtgtgacca aagtggctgtt gaagatgtt aagtcggacg
 2481 caacagagaa agacttgtca gacctgatct cagaaatggaa gatgatgaag atgatcggga
 2541 agcataagaa tatcatcaac ctgctggggg cctgcacgc ggtatggccc ttgtatgtca
 2601 tcgtggagta tgcctccaag ggcaacctgc gggagttactt gcaggccgg aggccccca
 2661 ggctggaata ctgctacaac cccagccaca acccagggaa gcagctctcc tccaaggacc
 2721 tgggtcctc cgcctaccag gtggcccgag gcatggagta tctggctcc aagaagtgc
 2781 tacaccgaga cctggcagcc aggaatgtcc tggtgacaga ggacaatgtg atgaagatag
 2841 cagactttgg cctcgacgg gacattcacc acatcgacta ctataaaaag acaaccaac
 2901 gcccactgcc tggtaagtgg atggcaccccg aggcatatt tgaccggatc tacacccacc
 2961 agagtgtatgtt gttggcttcc ggggtgcctc tggggagat cttcaactctg ggccggct
 3021 catacccccgg tggctgttg gaggaacttt tcaagctgtt gaaggagggtt caccggatgg
 3081 acaagcccaag taactgcacc aacgagctgtt acatgatgtt gggggactgc tggcatgc
 3141 tggccctaca gagaccacc tcaagcgc tggtgaaaga cctggaccgc atcggtggct
 3201 tgacccctcaa ccaggaggatc ctggacactgtt ccatggccctt ggaccaggatc tccccca
 3261 ttcccgacac ccggagctct acgtgtctt cagggggagaa ttccgttcc tctcatgac
 3321 cgctgcccga ggagccctgc ctgcggcgc acccageccaa gttgcaat ggcggactca
 3381 aacgccccctg a

The FGFR1, transcript variant 1 protein is a member of the FGFR family, where amino acid sequences are highly conserved between members and throughout evolution. FGFR family members differ from one another in their ligand affinities and tissue distribution. A full-length representative protein consists of an extracellular region, composed of three immunoglobulin-like domains, a single hydrophobic membrane-spanning segment, and a cytoplasmic tyrosine kinase domain. The extracellular portion of the protein interacts with fibroblast growth factors, setting in motion a cascade of downstream signals, ultimately influencing a myriad of biological processes including mitogen-

esis and differentiation. This particular family member binds both acidic and basic fibroblast growth factors and is involved in limb induction. Mutations in this gene have been associated with Pfeiffer syndrome, Jackson-Weiss syndrome, Antley-Bixler syndrome, osteoglophonic dysplasia, and autosomal dominant Kallmann syndrome. See, e.g., Dode et al., "Kallmann Syndrome: Fibroblast Growth Factor Signaling Insufficiency?" *J Mol Med* 82(11):725-34 (2004); Coumoul et al., "Roles of FGF Receptors in Mammalian Development and Congenital Diseases," *Birth Defects Res C Embryo Today* 69(4):286-304 (2003), which are hereby incorporated by reference in their entirety. Alternatively

119

spliced variants, which encode different protein isoforms, have been described; however, not all variants have been fully characterized.

The nucleic acid and amino acid sequences for FGFR1 variants 2-6 may be found using the following reference sequence ID numbers on GenBank: FGFR1, transcript variant 2 (GenBank Accession No. NM_015850), FGFR1, transcript variant 3 (GenBank Accession No. NM_023105), FGFR1, transcript variant 4 (GenBank Accession No. NM_023106), FGFR1, transcript variant 5 (GenBank Accession No. NM_023107), FGFR1, transcript variant 6 (GenBank Accession No. NM_023108), and FGFR1, transcript variant 9, (GenBank Accession No. NM_023111). These sequences are hereby incorporated by reference in their entirety.

Yet another aspect of the present invention relates to a method of causing increased FGF21 receptor agonist- β Klotho-FGFR complex formation. This method comprises providing a cell comprising β Klotho and an FGFR and providing an FGF21 receptor agonist, where the agonist comprises a chimeric protein comprising a C-terminal portion of FGF19. This method also includes contacting the cell and the FGF21 receptor agonist under conditions effective to cause increased FGF21 receptor agonist- β Klotho-FGFR complex formation relative to contacting the cell with FGF21 alone, where the FGF21 has a core domain.

With respect to the FGF21 agonist, suitable chimeric proteins include those chimeric proteins according to the present invention that are described above and throughout the present application. Suitable N-terminal portions of FGF21 and C-terminal portions of FGF19 are also described above and throughout the present application.

In one embodiment, the method of causing increased FGF21 receptor agonist- β Klotho-FGFR complex formation is carried out *in vitro*. In one embodiment, the method is carried out in an adipocyte.

In one embodiment, the method of causing increased FGF21 receptor agonist- β Klotho-FGFR complex formation is carried out *in vivo*. In one embodiment, the method is carried out in a mammal. In one particular embodiment, the mammal is a mouse.

A further aspect of the present invention relates to a method of screening for compounds with enhanced binding affinity for β Klotho suitable for fusion to the C-terminus of an N-terminal portion of FGF21 to generate an FGF21 agonist. The method includes providing FGF21, providing β Klotho, and providing one or more candidate compounds; combining the FGF21, the β Klotho, and the candidate compounds under conditions effective for FGF21 and β Klotho to form a binary complex if present by themselves; and identifying the candidate compounds which diminish binary complex formation, compared to when the candidate compound is absent, as being potentially suitable for fusion to the C-terminus of an N-terminal portion of FGF21 to generate an FGF21 agonist. In one embodiment, the candidate compound out-competes FGF21 for binding to the β Klotho.

Yet a further aspect of the present invention relates to a method of screening for compounds with enhanced binding affinity for the β Klotho-FGFR complex suitable for treatment of diabetes, obesity, or related metabolic disorders. This method includes providing FGF21, providing a binary β Klotho-FGFR complex, and providing one or more candidate compounds. This method also includes combining the FGF21, the binary β Klotho-FGFR complex, and the candidate compounds under conditions effective for the FGF21 and the β Klotho-FGFR complex to form a ternary complex

120

if present by themselves and identifying the candidate compounds which diminish ternary complex formation compared to when the candidate compound is absent as being potentially suitable for treatment of diabetes, obesity, or related metabolic disorders. In one embodiment, the candidate compound out-competes FGF21 for binding to the β Klotho-FGFR complex.

In one embodiment of the screening aspects of the present invention, the FGF21 has the amino acid sequence of SEQ ID NO: 100.

In one embodiment of the screening aspects of the present invention, β Klotho has the amino acid sequence of SEQ ID NO: 217 or SEQ ID NO: 218.

In one embodiment of the screening aspects of the present invention, the FGF receptor is FGFR1c. In one particular embodiment, the FGFR1c receptor has the amino acid sequence of SEQ ID NO: 221.

In one embodiment of the screening aspects of the present invention, a plurality of compounds is tested. In one embodiment, the candidate compounds are biomolecules. In one embodiment, the biomolecules are proteins. In one embodiment, the biomolecules are peptides. In one particular embodiment, the peptides are synthetic peptides. In one embodiment, the compounds are small organic molecules.

In one embodiment of the screening aspects of the present invention, the method is carried out using a cell-based assay. In one embodiment, the identifying is carried out using a cell-based assay.

In one embodiment of the screening aspects of the present invention, the method is carried out using a binding assay. In one embodiment, the binding assay is a direct binding assay. In one embodiment, the binding assay is a competition-binding assay. In one embodiment, the binding assay is carried out using surface plasmon resonance spectroscopy. In one embodiment, the identifying is carried out using a binding assay. In one embodiment, the identifying is carried out using surface plasmon resonance spectroscopy.

In one embodiment of the screening aspects of the present invention, the cell-based assay is carried out with adipocytes. In one embodiment, the cell-based assay is carried out with skeletal muscle cells. In one embodiment, stimulation of glucose uptake is the assay readout. In one embodiment, induction of glucose transporter 1 gene expression is the assay readout. In one embodiment, a dose-response curve is generated for the stimulation of glucose uptake by a candidate compound to determine potency and efficacy of the candidate compound. In one embodiment, a dose-response curve is generated for the induction of glucose transporter 1 gene expression by a candidate compound to determine potency and efficacy of the candidate compound. For example, if the dose-response curve is shifted to the left compared to that obtained for native FGF21, the candidate compound has greater potency than native FGF21. In one embodiment, an IC₅₀ value is derived from the dose-response curve of a candidate compound to determine potency of the candidate compound. An IC₅₀ value smaller than that obtained for native FGF21 identifies a candidate compound as more potent than native FGF21.

In one embodiment of the screening aspects of the present invention, the cell-based assay is carried out with mammalian cells ectopically expressing β Klotho. In one particular embodiment, the cells are HEK293 cells. In one embodiment, activation of FGF receptor is the assay readout. In one embodiment, tyrosine phosphorylation of an FGF receptor substrate is used as readout for FGF receptor activation. In one particular embodiment, the FGF receptor substrate is FGF receptor substrate 2 α . In one embodiment, activation of

121

downstream mediators of FGF signaling is used as readout for (or an indicator of) FGF receptor activation. In one particular embodiment, the downstream mediator of FGF signaling is 44/42 mitogen-activated protein kinase. In one embodiment, the downstream mediator of FGF signaling is a transcription factor. In one particular embodiment, the transcription factor is early growth response 1. In one embodiment, a dose-response curve is generated for β Klotho-dependent activation of FGF receptor by a candidate compound to determine potency and efficacy of the candidate compound. For example, if the dose-response curve is shifted to the left compared to that obtained for native FGF21, the candidate compound is more potent than native FGF21. In one embodiment, an IC_{50} value is derived from the dose-response curve of a candidate compound to determine potency of the candidate compound. An IC_{50} value smaller than that obtained for native FGF21 identifies a candidate compound as more potent than native FGF21.

In one embodiment of the screening aspects of the present invention, the surface plasmon resonance spectroscopy-based assay is carried out using FGF21 as ligand coupled to a biosensor chip. In one embodiment, mixtures of β Klotho ectodomain with increasing concentrations of a candidate compound are passed over a biosensor chip containing FGF21. In one embodiment, mixtures of the binary complex of FGFR ligand-binding domain and β Klotho ectodomain with increasing concentrations of a candidate compound are passed over a biosensor chip containing FGF21. In one particular embodiment, the FGFR ligand-binding domain is the FGFR1c ligand-binding domain. In one embodiment, an inhibition-binding curve is plotted for a candidate compound to determine potency of the candidate compound. For example, if the inhibition-binding curve is shifted to the left compared to that obtained for native FGF21, the candidate compound has greater potency than native FGF21. In one embodiment, an IC_{50} value is derived from the inhibition-binding curve of a candidate compound to determine potency of the candidate compound. An IC_{50} value smaller than that obtained for native FGF21 identifies a candidate compound as more potent than native FGF21. In one embodiment, the inhibition constant K_i is determined for a candidate compound to determine potency of the candidate compound. A K_i value smaller than that obtained for native FGF21 identifies a candidate compound as more potent than native FGF21.

In one embodiment of the screening aspects of the present invention, the method is carried out *in vivo*. In one embodiment, the method is carried out in a mammal. In one particular embodiment, the mammal is a mouse. In one embodiment, the ability of a candidate compound to potentiate the hypoglycemic effect of insulin is used as readout for FGF21-like metabolic activity. This involves fasting the mammal for a period of time prior to insulin injection and measuring fasting blood glucose levels. The mammal is then injected with insulin alone or co-injected with insulin plus a candidate compound. Blood glucose levels are measured at several time points after the injection. If a candidate compound potentiates the hypoglycemic effect of insulin to a greater degree than native FGF21 does, the candidate compound exhibits enhanced efficacy. Likewise, if a candidate compound potentiates the hypoglycemic effect of insulin to a similar degree than native FGF21 does but at a lower dose compared to that of FGF21 and/or for a longer period of time compared to FGF21, the candidate compound has enhanced agonistic properties.

122

EXAMPLES

The following examples are provided to illustrate embodiments of the present invention but are by no means intended to limit its scope.

Materials and Methods used in Examples 1-8

Purification of FGF19, FGF21, FGF23, FGFR, α Klotho, and β Klotho Proteins and Purification of FGF19, FGF21, and FGF23 Peptides

- 10 The three endocrine FGF ligands, and mutants, chimeras, and C-terminal peptides thereof, as well as the ligand-binding domain of FGFRs were expressed in *E. coli* BL21 (DE3) cells. The secreted, bioactive form of human FGF19 (R23 to K216 of SEQ ID NO: 1), human FGF21 (H29 to S209 of SEQ ID NO: 100), and human FGF23 (Y25 to I251 of SEQ ID NO: 223) was refolded *in vitro* from inclusion bodies, and purified by published protocols (Ibrahim et al., "Biochemical Analysis of Pathogenic Ligand-dependent FGFR2 Mutations Suggests Distinct Pathophysiological Mechanisms for Craniofacial and Limb Abnormalities," *Hum Mol Genet.* 13(19):2313-2324 (2004), Plotnikov et al., "Crystal Structures of Two FGF-FGFR Complexes Reveal the Determinants of Ligand-receptor Specificity," *Cell* 101 (4):413-424 (2000), which are hereby incorporated by reference in their entirety). In order to minimize proteolysis of FGF23, arginine residues 176 and 179 of the proteolytic cleavage site 176 RXXR 179 (with reference to SEQ ID NO: 223) were replaced with glutamine as it occurs in the phosphate wasting disorder "autosomal dominant hypophosphatemic rickets" (Anonymous, "Autosomal Dominant Hypophosphataemic Rickets is Associated with Mutations in FGF23," *Nat Genet* 26(3):345-348 (2000); White et al., "Autosomal-dominant Hypophosphatemic Rickets (ADHR) Mutations Stabilize FGF-23," *Kidney Int* 60(6):2079-2086 (2001), which are hereby incorporated by reference in their entirety).
- 15 20 25 30 35

Chimeras composed of a N-terminal portion of human FGF21 (H29 to V197, H29 to S190, or H29 to L167 of SEQ ID NO: 100) and a C-terminal portion of human FGF19 (T204 to K216, M197 to K216, or L169 to K216 of SEQ ID NO: 1), termed FGF21 $^{29-197}$ /FGF19 $^{204-216}$ (SEQ ID NO: 205), FGF21 $^{29-190}$ /FGF19 $^{197-216}$ (SEQ ID NO: 206), and FGF21 $^{29-167}$ /FGF19 $^{169-216}$ (SEQ ID NO: 207), respectively, were purified from inclusion bodies by the same protocol as the wild-type protein.

45 Likewise, two single mutants (Q104M and Y207F, SEQ ID NOS: 152 and 232, respectively) and one triple mutant (Y207F/A208E/S209K, SEQ ID NO: 233) of human FGF21 were purified by the same protocol as the wild-type protein. The C-terminal tail peptide of human FGF19 (M171 to K216 of SEQ ID NO: 1, termed FGF19 $^{C-tail}$) and the C-terminal tail peptide of human FGF21 (P168 to S209 of SEQ ID NO: 100, termed FGF21 $^{C-tail}$) were expressed as fusion peptides with a 50 residue-long N-terminal tag including a hexahistidine tag, and purified from the soluble cell lysate fraction by nickel affinity- and ion exchange chromatographies.

50 55 60 65 66 The N-terminally hexahistidine-tagged C-terminal tail peptide of human FGF23 (S180 to I251 of SEQ ID NO: 223, termed FGF23 $^{C-tail}$) was expressed and purified as described previously (Goetz et al., "Isolated C-terminal Tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *Proc Natl Acad Sci USA* 107(1):407-412 (2010), which is hereby incorporated by reference in its entirety).

A single mutant (M96T) of human FGF23 (SEQ ID NO: 224) was purified by the same protocol as the wild-type

123

protein. The proteolytic cleavage site $^{176}\text{RXXR}^{179}$ was not mutated in the M96T mutant protein. The wild-type FGF23 protein used as a control in the experiments with the M96T mutant also did not contain mutations at the proteolytic cleavage site.

Full-length human FGF homologous factor 1B (FHF1B; M1 to T181), which was used as a negative control for surface plasmon resonance (SPR) spectroscopy, was purified by a published protocol (Olsen et al., "Fibroblast Growth Factor (FGF) Homologous Factors Share Structural but not Functional Homology with FGFs," *J Biol Chem* 278(36):34226-34236 (2003), which is hereby incorporated by reference in its entirety).

The ligand-binding domain of each of the seven principal human FGFRs, namely FGFR1b (D142 to E374 of SEQ ID NO: 225), FGFR1c (D142 to R365 of SEQ ID NO: 221), FGFR2b (A140 to E366 of SEQ ID NO: 227), FGFR2c (N149 to E368 of SEQ ID NO: 226), FGFR3b (D147 to H358 of SEQ ID NO: 229), FGFR3c (D147 to E365 of SEQ ID NO: 228), and FGFR4 (Q144 to D355 of SEQ ID NO: 230), was refolded in vitro from inclusion bodies, and purified as described previously (Ibrahimi et al., "Biochemical Analysis of Pathogenic Ligand-dependent FGFR2 Mutations Suggests Distinct Pathophysiological Mechanisms for Craniofacial and Limb Abnormalities," *Hum Mol Genet* 13(19):2313-2324 (2004); Plotnikov et al., "Crystal Structures of Two FGF-FGFR Complexes Reveal the Determinants of Ligand-receptor Specificity," *Cell* 101(4):413-424 (2000), which are hereby incorporated by reference in their entirety).

The ectodomain of murine α Klotho (A35 to K982 of SEQ ID NO: 231) was purified from culture media of a HEK293 cell line ectopically expressing the α Klotho ectodomain as a fusion protein with a C-terminal FLAG tag (Kurosu et al., "Regulation of Fibroblast Growth Factor-23 Signaling by Klotho," *J Biol Chem* 281(10):6120-6123 (2006); Kurosu et al., "Suppression of Aging in Mice by the Hormone Klotho," *Science* 309(5742):1829-1833 (2005), which are hereby incorporated by reference in their entirety). Similarly, the ectodomain of murine β Klotho (F53 to L995 of SEQ ID NO: 218) was expressed in HEK293 cells as a fusion protein with a C-terminal FLAG tag and purified using the same protocol as for the α Klotho ectodomain. Purified bovine β -glucuronidase was obtained from Sigma-Aldrich.

Analysis of Ternary Complex Formation Between FGF19/FGF21, FGFR, and β Klotho by Size-Exclusion Chromatography

Size-exclusion chromatography experiments were performed on a HiLoad™ 16/60 Superdex™ 200 prep grade column (GE Healthcare) mounted on an ÄKTApurifier (GE Healthcare). Because of poor solubility of the ligand-binding domain of FGFR1c in low salt buffer, the experiments were carried out with 25 mM HEPES-NaOH buffer, pH 7.5, containing 1.0 M NaCl. Sample injection volume was 0.9 to 2.0 ml, and the flow rate was 1.0 ml min⁻¹. Protein retention times were determined by absorbance at 280 nm. The column was calibrated with ferritin (440 kDa), immunoglobulin G (150 kDa), albumin (69.3 kDa), ovalbumin (44.3 kDa), carbonic anhydrase (28.8 kDa), and ribonuclease A (13.7 kDa). The void volume was determined using blue dextran 2,000, and the column volume was measured with acetone. To examine ternary complex formation between FGF21, FGFR1c, and β Klotho, 2.72 μ mol of the 1:1 binary complex of FGFR1c ligand-binding domain and β Klotho ectodomain were mixed with 9.25 μ mol of FGF21, and the mixture was applied to the size-exclusion column. The retention time of the FGFR1c- β Klotho complex alone

124

served as a reference point. To examine ternary complex formation between FGF19, FGFR4, and β Klotho, 2.46 μ mol of the 1:1 binary complex of FGFR4 ligand-binding domain and β Klotho ectodomain were mixed with 8.51 μ mol of FGF19, and the mixture was applied to the size-exclusion column. The retention time of the FGFR4- β Klotho complex alone served as a reference point. Proteins of column peak fractions were resolved on 14% SDS-polyacrylamide gels, and then stained with Coomassie Brilliant Blue R-250.

10 Analysis of Mutant and Wild-Type FGF23 Proteins by Size-Exclusion Chromatography

Size-exclusion chromatography experiments were performed on a HiLoad™ 16/60 Superdex™ 75 prep grade column (GE Healthcare). Because of poor solubility of FGF23 in low salt buffer, the experiments were carried out with 25 mM HEPES-NaOH buffer, pH 7.5, containing 1.0 M NaCl. Sample injection volume was 1.5 to 3.8 ml, and the flow rate was 1.0 ml min⁻¹. Protein retention times were determined by absorbance at 280 nm. The column was calibrated with albumin (69.3 kDa), ovalbumin (44.3 kDa), carbonic anhydrase (28.8 kDa), ribonuclease A (13.7 kDa), and aprotinin (6.5 kDa). The void volume was determined using blue dextran 2,000, and the column volume was measured with acetone. To assess stability of FGF23 harboring the M96T mutation, equal amounts of mutant protein were injected onto the column at different times after affinity purification of the mutant protein. As a control, the elution profile of wild-type FGF23 was studied.

Analysis of FGF19/21/23- α / β Klotho, FGFR- β Klotho, and FGF21-FGFR1c- β Klotho Interactions by Surface Plasmon Resonance Spectroscopy

SPR experiments were performed on a Biacore 2000 instrument (Biacore AB), and all the protein-protein and protein-peptide interactions were studied at 25° C. in HBS-EP buffer (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% (v/v) polysorbate 20). Proteins were immobilized by amine coupling on flow channels of research grade CM5 chips (Biacore AB). Proteins were injected over a CM5 chip at a flow rate of 50 μ l min⁻¹, and at the end of each protein injection (180 s), HBS-EP buffer (50 μ l min⁻¹) was flowed over the chip to monitor dissociation for 180 s. In experiments where β Klotho binding to FGFR was analyzed, the chip surface was then regenerated by injecting 50 μ l of 2.0 M NaCl in 10 mM sodium/potassium phosphate, pH 6.5. In experiments where α / β Klotho binding to FGF19/21/23 or binding of the FGFR1c- β Klotho complex to FGF21 was studied, 2.0 M NaCl in 10 mM sodium acetate, pH 4.5 was used for chip regeneration. To control for nonspecific binding in experiments where the ectodomain of α Klotho or β Klotho was immobilized on the chip, β -glucuronidase was coupled to the control flow channel of the chip (~43-68 fmol/mm²). Like α Klotho and β Klotho, β -glucuronidase is a member of family 1 glycosidases (carbohydrate-active enzymes database at cazy's online website; Cantarel et al., "The Carbohydrate-Active EnZymes Database (CAZy): an Expert Resource for Glycogenomics," *Nucleic Acids Res* 37:D233-238 (2009), which is hereby incorporated by reference in its entirety), and hence structurally related to each of the two extracellular glycosidase-like domains of α Klotho and β Klotho, respectively. In experiments where an FGF ligand was immobilized on the chip, FHF1B, which shares structural similarity with FGFs, but does not exhibit any FGFR binding (Olsen et al., "Fibroblast Growth Factor (FGF) Homologous Factors Share Structural but not Functional Homology with FGFs," *J Biol Chem* 278(36):34226-34236 (2003), which is hereby incorporated by reference in its

125

entirety), was coupled to the control flow channel of the chip (~17-101 fmol/mm²). The data were processed with BiaEvaluation software (Biacore AB). For each protein injection over a chip onto which α Klotho or β Klotho had been immobilized, the nonspecific responses from the β -glucuronidase control flow channel were subtracted from the responses recorded for the α/β Klotho flow channel. Similarly, for each protein injection over a FGF chip, the non-specific responses from the FHF1B control flow channel were subtracted from the responses recorded for the FGF flow channel. Each set of experiments was repeated at least twice.

To analyze FGF21 binding to the binary FGFR1c- β Klotho complex, FGF21 was immobilized on a chip (~20 fmol/mm² of flow channel), and increasing concentrations of 1:1 complex of FGFR1c ligand-binding domain and β Klotho ectodomain in HBS-EP buffer were passed over the chip. To test the specificity of the interaction between FGF21 and the FGFR1c- β Klotho complex, two concentrations of 1:1 complex of FGFR1c ligand-binding domain and α Klotho ectodomain in HBS-EP buffer were passed over the FGF21 chip. The results are shown in FIGS. 1C and 1D.

To measure binding of β Klotho to each of the seven principal FGFRs, the ectodomain of β Klotho was immobilized on a chip (~42-46 fmol/mm² of flow channel). Increasing concentrations of the ligand-binding domain of FGFR1b, FGFR1c, FGFR2b, FGFR2c, FGFR3b, FGFR3c, or FGFR4 in HBS-EP buffer were passed over the chip. Maximal equilibrium binding responses were plotted against the concentrations of FGFR ligand-binding domain, and from the fitted saturation binding curve the equilibrium dissociation constant (K_D) was calculated. The fitted binding curve was judged to be accurate based on the distribution of the residuals (even and near zero) and χ^2 (<10% of R_{max}). The results are shown in FIGS. 3A to 3G.

To analyze binding of β Klotho to FGF19 and FGF21, FGF19 and FGF21 were coupled to two flow channels of a chip (~30 fmol/mm² of flow channel). As a control, FGF23 was also coupled to the chip (~29 fmol/mm² of flow channel). Increasing concentrations of the ectodomain of β Klotho in HBS-EP buffer were injected over the chip. As an additional control, binding of α Klotho to FGF19 and FGF21 was studied. The results are shown in FIGS. 4A to 4E.

To examine whether the isolated C-terminal tail peptide of FGF19 or FGF21 can compete with full-length FGF19 or FGF21 for binding to β Klotho, FGF19 and FGF21 were immobilized on two flow channels of a chip (~18-29 fmol/mm² of flow channel). Increasing concentrations of either FGF19^{C-tail} (0-20 nM) or FGF21^{C-tail} (0-200 nM) were mixed with a fixed concentration of β Klotho (10 nM) in HBS-EP buffer, and the mixtures were passed over the chip. To test the specificity of the interaction between β Klotho and the C-terminal tail of FGF19 or FGF21, β Klotho ectodomain was mixed with a 2-fold molar excess of FGF23^{C-tail}, and the mixture was injected over the chip. The results are shown in FIGS. 5B to 5G.

To examine whether mutants of FGF21 or chimeras composed of a N-terminal portion of FGF21 and a C-terminal portion of FGF19 can compete with wild-type ligand for binding to the FGFR1c- β Klotho complex, FGF21 was immobilized on a chip (~30 fmol/mm² of flow channel). Increasing concentrations of FGF21 mutant or chimera (0-60 nM) were mixed with a fixed concentration of 1:1 complex of FGFR1c ligand-binding domain and β Klotho ectodomain (10 nM), and the mixtures were passed over the chip. As a control, competition of FGF21 in solution with

126

immobilized FGF21 for binding to the FGFR1c- β Klotho complex was studied. The results are shown in FIGS. 14A to 14D and 15A to 15F.

To examine whether a chimera composed of a N-terminal portion of FGF21 and a C-terminal portion of FGF19 can compete with wild-type FGF21 for binding to β Klotho, FGF21 was immobilized on a chip (~29 fmol/mm² of flow channel). Increasing concentrations of chimera (0-40 nM) were mixed with a fixed concentration of β Klotho ectodomain (10 nM), and the mixtures were injected over the chip. As a control, competition of FGF21 in solution with immobilized FGF21 for binding to β Klotho was studied. The results are shown in FIGS. 7A to 7C.

Analysis of Phosphorylation of FRS2 α and 44/42 MAP Kinase in a Hepatoma Cell Line

To test whether the C-terminal tail peptides of FGF19 and FGF21 are interchangeable in inhibiting the signaling of FGF19, H4IIE rat hepatoma cells, which endogenously express β Klotho and FGFR4, were serum starved overnight and then pretreated for 60 min with either FGF19^{C-tail} (10 to 1000 ng ml⁻¹) or FGF21^{C-tail} (10 to 1000 ng ml⁻¹) prior to stimulation with FGF19 (30 ng ml⁻¹) for 10 min. Cell stimulation with FGF19 (3 to 300 ng ml⁻¹), FGF19^{C-tail} (10 to 1000 ng ml⁻¹), or FGF21^{C-tail} (10 to 1000 ng ml⁻¹) alone served as controls.

After stimulation, the cells were lysed (Kurosu et al, "Suppression of Aging in Mice by the Hormone Klotho," *Science* 309(5742):1829-1833 (2005), which is hereby incorporated by reference in its entirety), and cellular proteins were resolved on SDS-polyacrylamide gels and transferred to nitrocellulose membranes. The protein blots were probed with an antibody to phosphorylated FGFR substrate-2 α (FRS2 α), and with antibodies recognizing only phosphorylated 44/42 MAP kinase or both phosphorylated and nonphosphorylated (total) 44/42 MAP kinase. All antibodies were from Cell Signaling Technology. The results are shown in FIGS. 6A and 6B.

Analysis of Egr1 Protein Expression in an Epithelial Cell Line

The ability of a single mutant of FGF21 and an FGF21/FGF19 chimera to activate FGFR1c in a β Klotho-dependent fashion was studied using induction of Egr1 expression as readout for FGFR1c activation. HEK293 cells, which endogenously express FGFR1c (Kurosu et al, "Regulation of Fibroblast Growth Factor-23 Signaling by Klotho," *J Biol Chem* 281:6120-6123 (2006), which is hereby incorporated by reference in its entirety), were transiently transfected with murine β Klotho. β Klotho HEK293 transfectants were serum starved overnight and then stimulated for 90 min with FGF21 mutant, chimera, or wild-type protein (3 to 300 ng ml⁻¹ each). After stimulation, the cells were lysed (Kurosu et al, "Suppression of Aging in Mice by the Hormone Klotho," *Science* 309(5742):1829-1833 (2005), which is hereby incorporated by reference in its entirety), and cellular proteins were resolved on SDS-polyacrylamide gels and transferred to nitrocellulose membranes. The protein blots were probed with antibodies to Egr1 and GAPDH. The intensity of the protein bands on the immunoblots was quantified and the ratio of Egr1 to GAPDH was calculated. The ratio of Egr1 to GAPDH was then plotted as a function of FGF21 ligand concentration. The anti-Egr1 antibody was from Cell Signaling Technology and the anti-GAPDH antibody was from Abcam. The results are shown in FIGS. 17A to 17C.

Insulin Tolerance Test in Mice

The metabolic activity of a single mutant of FGF21 and an FGF21/FGF19 chimera was studied in C57BL/6 mice.

127

The ability of FGF21 mutant or chimera to potentiate the hypoglycemic effect of insulin was used as readout for FGF21-like metabolic activity (Ohnishi et al., "Dietary and Genetic Evidence for Enhancing Glucose Metabolism and Reducing Obesity by Inhibiting Klotho Functions," *FASEB J* 25, 2031-2039 (2011), which is hereby incorporated by reference in its entirety). Mice were kept on normal chow. On the day of the insulin tolerance test, mice were fasted for 4 h and then bled from the cheek pouch for measuring fasting blood glucose levels. Thereafter, mice were administered intraperitoneally insulin (0.5 units per kilogram body weight) alone or insulin (0.5 units per kilogram body weight) plus either FGF21 mutant or FGF21/FGF19 chimera (0.3 mg per kilogram body weight). As controls, mice were injected with vehicle alone or co-injected with insulin plus FGF21. At the indicated time points after the injection (FIGS. 18A-18C), blood was drawn from the tail vein. Glucose concentrations in the blood samples were determined using Bayer Contour® blood glucose test strips (Bayer Corp.).

Example 1

Klotho Co-Receptors Use Different Mechanisms to Promote Binding of Endocrine FGF Ligands to Cognate FGFRs

The protein-protein interactions leading to the formation of the ternary complex between FGF23, FGFR1c, and α Klotho were previously characterized (Goetz et al., "Isolated C-terminal Tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *Proc Natl Acad Sci USA* 107(1):407-412 (2010), which is hereby incorporated by reference in its entirety). It was shown that the ectodomain of α Klotho possesses a high-affinity binding site for the ligand-binding domain of FGFR1c but not for the FGF23 ligand (Goetz et al., "Isolated C-terminal Tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *Proc Natl Acad Sci USA* 107(1):407-412 (2010), which is hereby incorporated by reference in its entirety), and that the preformed binary FGFR1c- α Klotho complex binds avidly to FGF23 (Goetz et al., "Isolated C-terminal Tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *Proc Natl Acad Sci USA* 107(1):407-412 (2010), which is hereby incorporated by reference in its entirety). It was concluded that FGF23 binds to a de novo binding site generated at the composite FGFR1c- α Klotho interface. The region on FGF23 that binds to this site was mapped to the C-terminal tail that follows the β -trefoil core domain (Goetz et al., "Isolated C-terminal Tail of FGF23 Alleviates Hypophosphatemia by Inhibiting FGF23-FGFR-Klotho Complex Formation," *Proc Natl Acad Sci USA* 107(1):407-412 (2010), which is hereby incorporated by reference in its entirety). Here it was explored whether β Klotho uses the same mechanism to promote binding of FGF19 and FGF21 to FGFR4 and FGFR1c, the principal cognate FGFRs of these ligands. It was first examined whether the FGF21-FGFR1c- β Klotho ternary complex can be reconstituted in solution in the same manner as the FGF23-FGFR1c- α Klotho complex. To form FGFR1c- β Klotho binary complex, conditioned media from a HEK293 cell line ectopically expressing murine β Klotho ectodomain (F53 to L995 of SEQ ID NO: 218) was applied to an affinity column containing the ligand-binding domain of FGFR1c (D142 to R365 of SEQ ID NO: 221). The FGFR1c- β Klotho complex eluted from the column was

128

purified further by size-exclusion chromatography (FIG. 1A). To examine ternary complex formation, the FGFR1c- β Klotho complex was mixed with FGF21 (H29 to S209 of SEQ ID NO: 100), and the mixture was applied to a size-exclusion column. As shown in FIG. 1B, FGF21 coeluted with FGFR1c- β Klotho, demonstrating that, similar to FGF23, FGF21 forms a stable ternary complex with the ectodomain of β Klotho and the ligand-binding domain of FGFR1c. Consistent with the gel filtration data, analysis of 5 ternary complex formation by SPR spectroscopy also showed that FGF21 binds the binary FGFR1c- β Klotho complex (FIG. 1C). The SPR analysis further showed that FGF21 does not interact with the FGFR1c- α Klotho complex demonstrating that the interaction between FGF21 and the 10 FGFR1c- β Klotho complex is specific (FIG. 1D). Similar to FGF21, the ternary complex of FGF19 with its cognate receptor (FGFR4) and β Klotho co-receptor could be reconstituted in solution using the ectodomain of β Klotho and the 15 ligand-binding domain of FGFR4 (FIGS. 2A and 2B).

Since the ectodomain of β Klotho forms stable binary complexes with the ligand-binding domains of FGFR1c and FGFR4, it was reasoned that it must contain a high affinity 20 binding site for FGFR1c and FGFR4. To substantiate this and to measure the binding affinity of β Klotho for each of 25 the two receptors, SPR spectroscopy was employed. β Klotho ectodomain was immobilized on a biosensor chip, and increasing concentrations of the ligand-binding domain of either FGFR1c or FGFR4 were passed over the chip. β Klotho bound both receptors with comparably high affinity 30 (FIGS. 3A and 3D), demonstrating that similar to α Klotho, β Klotho contains a high-affinity binding site for its cognate FGFRs.

For ternary complex formation with FGF19 or FGF21, two possible mechanisms remained open: one was that a de 35 novo binding site for the ligand was generated in the context of the binary β Klotho-FGFR complex as in the case of ternary complex formation between α Klotho, FGF23, and FGFR; the other possibility was that β Klotho contained a distinct high affinity binding site for the ligand. In order to 40 distinguish between these two mechanisms, it was examined, by SPR spectroscopy, whether β Klotho directly binds to FGF 19 and FGF21, respectively. FGF19 and FGF21 and as a specificity control, FGF23 were immobilized on a biosensor chip, and increasing concentrations of the ectodomain of β Klotho were passed over the chip. Both FGF19 and FGF21 bound strongly to β Klotho (FIGS. 4A and 4B), whereas no interaction was observed between FGF23 and β Klotho (FIG. 4C). To further confirm the specificity of the 45 interaction, increasing concentrations of the ectodomain of α Klotho were passed over the chip. Neither FGF19 nor FGF21 bound to α Klotho (FIGS. 4D and 4E). Together, the data show that in contrast to α Klotho, β Klotho possesses distinct high-affinity binding sites for cognate endocrine FGF ligand and FGFR, indicating that β Klotho promotes 50 ternary complex formation by engaging FGF ligand and FGFR simultaneously.

Example 2

β Klotho Binding Site on FGF19 and FGF21 Maps to the C-Terminal Region of Each Ligand

It was next investigated which sequences of FGF19 and FGF21 bind to β Klotho. A clue to the location of the 60 β Klotho binding site on FGF19 and FGF21 came from the previous finding that the binding site on FGF23 for the binary FGFR- α Klotho complex resides in the C-terminal 65

129

region of FGF23 that follows the β -trefoil core domain (Goetz et al., "Molecular Insights into the Klotho-Dependent, Endocrine Mode of Action of Fibroblast Growth Factor 19 Subfamily Members," *Mol Cell Biol* 27(9):3417-3428 (2007), which is hereby incorporated by reference in its entirety). Subsequent studies suggested that the same region in FGF19 and FGF21 mediates binding of these ligands to β Klotho. Specifically, it was shown that a chimera of FGF19 with the C-terminal tail of FGF21 was able to bind β Klotho and gradual deletion of C-terminal residues of FGF21 resulted in progressively reduced binding affinity for β Klotho (Wu et al., "C-terminal Tail of FGF19 Determines its Specificity Toward Klotho Co-receptors," *J Biol Chem* 283(48):33304-33309 (2008); Yie et al., "FGF21 N- and C-termini Play Different Roles in Receptor Interaction and Activation," *FEBS Lett* 583(1):19-24 (2009); Micanovic et al., "Different Roles of N- and C-termini in the Functional Activity of FGF21," *J Cell Physiol* 219(2):227-234 (2009), which are hereby incorporated by reference in their entirety). In order to unambiguously demonstrate that the β Klotho-binding site on FGF19 and FGF21 resides in the C-terminal region of each ligand, the C-terminal tail peptides of FGF19 (FGF19^{C-tail}; M171 to K216 of SEQ ID NO: 1) and FGF21 (FGF21^{C-tail}; P168 to S209 of SEQ ID NO: 100) were expressed and purified. It was then examined, by SPR spectroscopy, whether each peptide can compete with full-length ligand for binding to β Klotho. FGF19 and FGF21 were immobilized on a biosensor chip, and mixtures of a fixed concentration of β Klotho ectodomain with increasing concentrations of either FGF19^{C-tail} or FGF21^{C-tail} were passed over the chip. As shown in FIG. 5B, FGF19^{C-tail} competed, in a dose-dependent fashion, with FGF19 for binding to β Klotho. Similarly, FGF21^{C-tail} competed with FGF21 for binding to β Klotho (FIG. 5E). To confirm that the interaction between β Klotho and the C-terminal tail of FGF19 or FGF21 is specific, β Klotho ectodomain was mixed with a 2-fold molar excess of the C-terminal tail peptide of FGF23 (FGF23^{C-tail}), and the mixture was passed over the FGF19/21 chip. As expected, FGF23^{C-tail} did not interfere with β Klotho binding to immobilized FGF19 or FGF21 (FIGS. 5D and 5G). Together, the data conclusively show that the C-terminal region of FGF19 and FGF21 contains the β Klotho-binding site.

Example 3

FGF19 and FGF21 Share a Common Binding Site on β Klotho

Since both FGF19 and FGF21 bind to β Klotho, it raised the question whether these ligands bind to a shared site on β Klotho or whether each ligand has its own distinct binding site. To answer this, an SPR-based competition binding assay as described above was employed to examine whether the isolated C-terminal tail peptide of FGF19 can compete with full-length FGF21 for binding to β Klotho, and conversely, whether the C-terminal tail peptide of FGF21 can compete with full-length FGF19 for binding to β Klotho. As shown in FIG. 5F, FGF19^{C-tail} effectively competed with FGF21 for binding to β Klotho. Similarly, FGF21^{C-tail} was capable of inhibiting β Klotho binding to FGF19 (FIG. 5C). These data show that FGF19 and FGF21 have overlapping binding sites on β Klotho.

To provide biological evidence for the in vitro finding that FGF19 and FGF21 bind to a shared binding site on β Klotho, it was next examined whether the FGF19^{C-tail} peptide and the FGF21^{C-tail} peptide are both able to block FGF19 sig-

130

naling in cells. H4IE hepatoma cells, which endogenously express β Klotho and FGFR4 (Kurosu et al., "Tissue-specific Expression of betaKlotho and Fibroblast Growth Factor (FGF) Receptor Isoforms Determines Metabolic Activity of FGF19 and FGF21," *J Biol Chem* 282:26687-26695 (2007), which is hereby incorporated by reference in its entirety), were pretreated with FGF19^{C-tail} or FGF21^{C-tail} and then stimulated with FGF19. As shown in FIGS. 6A and 6B, both FGF19^{C-tail} and FGF21^{C-tail} inhibited, in a dose-dependent fashion, FGF19-induced tyrosine phosphorylation of FRS2a and downstream activation of MAP kinase cascade. As expected, neither of the two peptides elicited any signaling response when applied alone (FIGS. 6A and 6B). These data show that the C-terminal tail peptides of FGF19 and FGF21 are interchangeable in inhibiting the signaling of FGF19, and provide cell-based evidence that FGF19 and FGF21 share a common binding site on β Klotho. Importantly, the binding site overlap may provide a molecular mechanism for why transgenic expression or therapeutic administration of FGF 19 produces beneficial effects on glucose and lipid metabolism resembling those elicited by FGF21 (Fu et al., "Fibroblast Growth Factor 19 Increases Metabolic Rate and Reverses Dietary and Leptin-deficient Diabetes," *Endocrinology* 145:2594-2603 (2004); Tomlinson et al., "Transgenic Mice Expressing Human Fibroblast Growth Factor-19 Display Increased Metabolic Rate and Decreased Adiposity," *Endocrinology* 143:1741-1747 (2002), which are hereby incorporated by reference in their entirety).

Example 4

FGF19 Binds β Klotho with Greater Affinity than FGF21

It was next asked whether FGF19 and FGF21 bind with similar affinity to the common site on β Klotho or whether the two ligands have different binding affinities for β Klotho. A quantitative analysis of the SPR data shows that the FGF 19 C-terminal tail peptide is more potent than the FGF21 C-terminal tail peptide at inhibiting binding of β Klotho to full-length FGF 19 or FGF21. Specifically, an equimolar amount of FGF19^{C-tail} relative to β Klotho already yielded nearly complete inhibition of β Klotho binding to FGF19 or FGF21 (FIGS. 5B and 5F), whereas a 10- to 20-fold molar excess of FGF21^{C-tail} over β Klotho was needed to achieve a similar effect (FIGS. 5C and 5E). These data indicate that the C-terminal tail of FGF19 binds β Klotho with greater affinity than the C-terminal tail of FGF21 suggesting that primary sequence differences at this region account for the observed difference in binding affinity of the two ligands for β Klotho.

Comparison of the C-terminal tail sequences of FGF19 and FGF21 shows a significant degree of sequence similarity (40% amino acid identity) only in the last twenty residues (FIG. 5A; see also FIG. 8B), pointing to these residues as the major binding epitope for β Klotho. To test this possibility, the nineteen most C-terminal residues in FGF21 were swapped with the corresponding residues of FGF19, including a one-residue insertion (FIG. 12, FGF21^{C-tail} variant 19-45), and it was examined, by SPR spectroscopy, whether the chimeric FGF21 protein (termed FGF21¹²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶; SEQ ID NO: 206) is more potent than wild-type FGF21 at inhibiting binding of β Klotho to immobilized FGF21. As shown in FIGS. 7B and 7C, an equimolar amount of FGF21¹²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera relative to β Klotho yielded nearly complete inhibition of β Klotho binding to immobilized FGF21, whereas the same molar ratio of wild-type FGF21 to β Klotho produced at best half-maximum

131

inhibition (FIGS. 7A and 7C). These data show that the exchange of unique residues in the distal portion of the C-terminal tail of FGF21 for the corresponding residues of FGF 19 confers increased binding affinity to β Klotho on FGF21. In other words, the sequence from M197 to K216 of FGF19 contains residues that contribute to the higher β Klotho-binding affinity of FGF19 compared to FGF21.

The biological significance of the differential binding affinities of FGF19 and FGF21 for β Klotho was next explored. Since FGF19 binds β Klotho with greater affinity than FGF21 does, FGF19 would out-compete FGF21 for β Klotho if both FGF ligands were present in target tissue at the same time. Under physiological conditions, FGF19 and FGF21 do not appear to equally coexist in the blood circulation (Badman et al., "Hepatic Fibroblast Growth Factor 21 is Regulated by PPARalpha and is a Key Mediator of Hepatic Lipid Metabolism in Ketotic States," *Cell Metab* 5:426-437 (2007); Galman et al., "The Circulating Metabolic Regulator FGF21 is Induced by Prolonged Fasting and PPARalpha Activation in Man," *Cell Metab* 8:169-174 (2008); Holt et al., "Definition of a Novel Growth Factor-dependent Signal Cascade for the Suppression of Bile Acid Biosynthesis," *Genes Dev* 17:1581-1591 (2003); Inagaki et al., "Fibroblast Growth Factor 15 Functions as an Enterohepatic Signal to Regulate Bile Acid Homeostasis," *Cell Metab* 2:217-225 (2005); Inagaki et al., "Endocrine Regulation of the Fasting Response by PPARalpha-mediated Induction of Fibroblast Growth Factor 21," *Cell Metab* 5:415-425 (2007); Tong et al., "Transcriptional Repressor E4-binding Protein 4 (E4BP4) Regulates Metabolic Hormone Fibroblast Growth Factor 21 (FGF21) During Circadian Cycles and Feeding," *J Biol Chem* 285:36401-36409 (2010), which are hereby incorporated by reference in their entirety). It was speculated that the high affinity interaction between FGF19 and β Klotho, together with the binding preference of β Klotho for FGFR4, ensure that most of the postprandially secreted FGF19 acts on the liver (and the gall bladder) and hence becomes trapped in the enterohepatic circulation. Importantly, these findings have provided for the rational design of an FGF21 agonist, as follows.

Example 5

Chimera Composed of a N-terminal Portion of FGF21 and a C-Terminal Portion of FGF19 Exhibits Enhanced Binding Affinity for the FGFR1c- β Klotho Complex

Based on these findings, it was reasoned that variants of FGF21 in which C-terminal residues unique to FGF21 were replaced with the corresponding residues of FGF19 should have enhanced binding affinity for β Klotho compared to native FGF21, and hence agonist potency. To begin to explore this, residues located in the distal portion of the C-terminal tail of FGF21 were progressively mutated, namely residues within the sequence from S191 to S209, since this region is essential in determining the ligand's binding affinity for β Klotho (FIG. 7). Specifically, a single mutant of FGF21 (Y207F; FIG. 11, FGF21^{C-tail} variant 19-3), a triple mutant of FGF21 (Y207F/A208E/S209K; FIG. 12, FGF21^{C-tail} variant 19-36), and a chimeric FGF21 protein in which the twelve most C-terminal residues in FGF21 were swapped with the corresponding residues of FGF19, including a one-residue insertion, (termed FGF21²⁹⁻¹⁹⁷/FGF19²⁰⁴⁻²¹⁶; SEQ ID NO: 205; FIG. 12, FGF21^{C-tail} variant 19-41) were made. The FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera (SEQ ID NO: 206; see FIG. 12,

132

FGF21^{C-tail} variant 19-45), which is discussed above, was also included in these studies. In this chimera, the entire sequence from S191 to S209 of FGF21 is replaced by the corresponding sequence of FGF19 (FIG. 12, FGF21^{C-tail} variant 19-45), and it was shown that this chimera exhibits enhanced binding affinity for β Klotho compared to native FGF21 (FIG. 7). A chimera in which the entire C-terminal tail of FGF21 was exchanged for the corresponding region of FGF 19 (termed FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶, SEQ ID NO: 207; FIG. 8A) was used as a control.

To test whether the FGF21 mutant or chimeric proteins exhibit agonist potency, a SPR-based competition binding assay was employed. A competition binding assay was selected over a direct binding assay because its binding data are not confounded by the effects that the coupling of one binding partner to the chip might have. Specifically, it was examined whether a mutant or chimera can compete with native FGF21 for binding to the FGFR1c- β Klotho complex. If a mutant or chimera had greater affinity for the FGFR1c- β Klotho complex than native FGF21, and hence agonist potency, it would out-compete native FGF21 for binding to FGFR1c- β Klotho. FGF21 was immobilized on a biosensor chip, and mixtures of a fixed concentration of FGFR1c- β Klotho complex with increasing concentrations of either FGF21 mutant or FGF21/FGF19 chimera were passed over the chip. As a control, competition of FGF21 in solution with immobilized FGF21 for binding to the FGFR1c- β Klotho complex was studied.

As expected, FGF21 in solution competed, in a dose-dependent fashion, with immobilized FGF21 for binding to the FGFR1c- β Klotho complex (FIGS. 14A and 15A). The Y207F mutant of FGF21 was a weaker competitor than wild-type FGF21 for binding to the FGFR1c- β Klotho complex (FIG. 14B), suggesting that the mutant had reduced affinity for FGFR1c- β Klotho compared to wild-type FGF21. The Y207F/A208E/S209K triple mutant of FGF21 was even less potent than the Y207F single mutant at inhibiting binding of the FGFR1c- β Klotho complex to immobilized wild-type FGF21 (FIGS. 14C and 14D). These data indicate that the triple mutation causes an even greater loss in binding affinity of FGF21 for FGFR1c- β Klotho than the single mutation.

Based on these findings, it was concluded that replacing Y207 in FGF21 with phenylalanine of FGF19 reduces rather than enhances the binding affinity of FGF21 for β Klotho, and the combined replacement of Y207, A208, and S209 for the corresponding residues of FGF19 has an even greater negative impact on the binding affinity of FGF21 for β Klotho. In contrast to those two mutants of FGF21, all three FGF21/FGF19 chimeras proved to be more potent competitors than native FGF21 for binding to the FGFR1c- β Klotho complex (FIGS. 15A-15F). At any given concentration tested, the FGF21²⁹⁻¹⁹⁷/FGF19²⁰⁴⁻²¹⁶ chimera caused greater inhibition of FGFR1c- β Klotho binding to the FGF21 chip surface than native FGF21 did (FIGS. 15A, 15B, and 15D), suggesting that it has increased affinity for FGFR1c- β Klotho compared to native FGF21. The FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera, which harbors four additional amino acid substitutions in the distal portion of the C-terminal tail of FGF21 compared to the FGF21²⁹⁻¹⁹⁷/FGF19²⁰⁴⁻²¹⁶ chimera (FIG. 12), exhibited a further increased binding affinity for FGFR1c- β Klotho; at any given dose, it inhibited binding of FGFR1c- β Klotho to immobilized FGF21 to a substantially greater degree than the FGF21²⁹⁻¹⁹⁷/FGF19²⁰⁴⁻²¹⁶ chimera (FIGS. 15E and 15F). For example, an equimolar amount of FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera relative to FGFR1c- β Klotho complex

133

yielded nearly complete inhibition of FGFR1c- β Klotho binding to immobilized FGF21, whereas the same molar ratio of FGF21²⁹⁻¹⁹⁷/FGF19²⁰⁴⁻²¹⁶ chimera to FGFR1c- β Klotho produced at best 75% inhibition (FIG. 15F). These data show that a substantial further increase in binding affinity for FGFR1c- β Klotho was achieved by introducing four amino acid substitutions in the FGF21 sequence from S191 to V197 in addition to replacing the unique residues C-terminal to V197 with the analogous residues of FGF19. The FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera in which the entire C-terminal tail of FGF21 was exchanged for the corresponding region of FGF19 was only slightly more potent than the FGF21²⁹⁻¹⁹⁰/FGF19¹⁹⁷⁻²¹⁶ chimera at inhibiting binding of FGFR1c- β Klotho to immobilized FGF21 (FIGS. 15E and 15F). Thus, the major increase in binding affinity for FGFR1c- β Klotho was obtained by replacing the C-terminal sequence from S191 to S209 in FGF21 with the analogous sequence of FGF19.

Together, the data show that an FGF21 agonist can be engineered by replacing C-terminal sequences in FGF21 with the corresponding sequences of FGF19. Increased binding affinity for β Klotho underlies the agonist potency of an FGF21/FGF19 chimera. Based on the findings with the triple mutant of FGF21, it was speculated that replacing the sequence from S191 to S206 in FGF21 with the analogous sequence of FGF19 might be sufficient to confer similar binding affinity for β Klotho on FGF21 as FGF19 has. Moreover, it is thought that replacing poorly conserved residues in the C-terminal region of FGF19 might further enhance the binding affinity of FGF19 itself for β Klotho (FIGS. 10 and 13).

Example 6

Chimera Composed of a N-Terminal Portion of FGF21 and a C-Terminal Portion of FGF19 Acts as an FGF21 Agonist in a Cell-Based Assay

The FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera, which has proved the most potent among the three FGF21/FGF19 chimeras in the competition binding experiments, was then selected for analysis of agonist potency and efficacy in a cell-based assay. Specifically, the ability of the chimera to activate FGFR1c in a β Klotho-dependent fashion in HEK293 cells co-expressing FGFR1c and β Klotho was examined. Induction of protein expression of Egr1, a known downstream mediator of FGF signaling, was used as readout for FGFR1c activation. As shown in FIG. 17A, the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera induced, in a dose-dependent fashion, Egr1 protein expression. The effect became evident at a 10-fold lower concentration of chimera than native FGF21 (FIG. 17A). The dose-response curve for the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera obtained from quantitative analysis of the data shown in FIG. 17A was markedly shifted to the left compared to the dose-response curve for native FGF21 (FIG. 17B). The maximum signaling responses were similar, however. These data show that the FGF21/FGF19 chimera exhibits greater potency than native FGF21, which is consistent with the SPR results.

Example 7

Chimera Composed of a N-Terminal Portion of FGF21 and a C-Terminal Portion of FGF19 Acts as an FGF21 Agonist In Vivo

These findings prompted examination of whether the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera exhibits FGF21 agonist

134

activity in vivo. Specifically, insulin tolerance was used as pharmacodynamic marker, and it was analyzed whether the chimera can potentiate the hypoglycemic effect of exogenous insulin in mice. As shown in FIGS. 18A and 18B, the FGF21²⁹⁻¹⁶⁷/FGF19¹⁶⁹⁻²¹⁶ chimera increased the hypoglycemic effect of insulin to a similar degree as native FGF21 did. However, the effect of the chimera persisted for at least twice as long as that of native FGF21 (FIGS. 18A and 18B). These data show that compared to native FGF21, the FGF21/FGF19 chimera has a prolonged potentiating effect on insulin-induced hypoglycemia, which is indicative of agonist potency.

Example 8

Mutant FGF21 Harboring Q104M Substitution in the Core Domain Acts as an FGF21 Agonist IN VITRO and In Vivo

In a second approach of engineering an FGF21 agonist, glutamine at position 104 in FGF21 was mutated to methionine in order to increase the thermal stability of the β -trefoil core domain of FGF21 (FGF21^{Q104M}, SEQ ID NO: 152). Except for FGF21, all FGF ligands have a methionine residue at the position analogous to Q104 of FGF21 (Mohammadi et al., "Structural Basis for Fibroblast Growth Factor Receptor Activation," *Cytokine & Growth Factor Rev* 16(2):107-137 (2005), which is hereby incorporated by reference in its entirety). Together with other hydrophobic residues, the methionine forms the interior hydrophobic core of an FGF ligand's β -trefoil core domain. The key role the methionine plays in providing stabilizing interactions in the hydrophobic core is evidenced by the fact that its replacement with threonine as it naturally occurs in FGF23 dramatically reduces protein stability (FIGS. 16A-E), and leads to disease (Chefetz et al., "A Novel Homozygous Missense Mutation in FGF23 Causes Familial Tumoral Calcinosis Associated with Disseminated Visceral Calcification," *Hum Genet.* 118(2):261-266 (2005), which is hereby incorporated by reference in its entirety).

In particular, as shown in FIGS. 16A-E, substitution of methionine at position 96 for threonine in FGF23 (SEQ ID NO: 224), as it occurs in Familial Tumoral Calcinosis (Chefetz et al., "A Novel Homozygous Missense Mutation in FGF23 Causes Familial Tumoral Calcinosis Associated with Disseminated Visceral Calcification," *Hum Genet.* 118 (2):261-266 (2005), which is hereby incorporated by reference in its entirety), destabilizes the FGF23 protein. FIG. 16A shows a molecular surface representation of the FGF23 crystal structure (PDB ID: 2P39; Goetz et al., "Molecular Insights into the Klotho-Dependent, Endocrine Mode of Action of Fibroblast Growth Factor 19 Subfamily Members," *Mol Cell Biol* 27:3417-3428 (2007), which is hereby incorporated by reference in its entirety). A close-up view into the hydrophobic interior core of FGF23's β -trefoil core domain showing some of the key hydrophobic side chains is shown on the right, and a view of the whole structure is shown on the left. Note that M96 makes numerous hydrophobic contacts with its neighboring residues such as I102, F115, and V136 in the β -trefoil core of FGF23. The M96T substitution would weaken these hydrophobic contacts leading to thermal instability of the FGF23 protein. FIG. 16B shows a size-exclusion chromatogram of the M96T mutant of FGF23 analyzed immediately after Ni-chelating affinity purification. FIG. 16C shows a size-exclusion chromatogram of the M96T mutant of FGF23 analyzed following incubation at 4° C. for 24 hours. FIG. 16D shows a size-

135

exclusion chromatogram of wild-type FGF23 immediately following protein purification. FIG. 16E shows a size-exclusion chromatogram of purified wild-type FGF23 following incubation at 4° C. for 24 hours. Arrows in FIGS. 16B-E indicate the retention times of molecular size standards, the void volume (V_v) and the column volume (V_c). Note that, in contrast to wild-type FGF23, there is a substantial increase in the portion of M96T mutant protein eluting in the void volume indicating that the mutant protein unfolds over time.

Thus, it was reasoned that substituting Q104 of FGF21 for methionine would confer greater stability on FGF21, and hence increase the half-life of the FGF21 protein in the blood circulation. Owing to its increased half-life compared to wild-type FGF21, the Q104M mutant might exhibit agonist potency.

To test this, a cell-based assay was first employed. Specifically, it was analyzed whether the mutant protein can activate FGFR1c in a β Klotho-dependent fashion in HEK293 cells co-expressing FGFR1c and β Klotho. Induction of protein expression of Egr1, a known downstream mediator of FGF signaling, was used as readout for FGFR1c activation. As shown in FIG. 17A, the Q104M mutant of FGF21 induced, in a dose-dependent fashion, Egr1 protein expression. The induction of Egr1 protein expression by the FGF21 mutant was already detectable at a concentration of 30 ng ml⁻¹, whereas a more than 3-fold greater concentration of wild-type FGF21 was needed to see a similar effect

136

(FIG. 17A). The dose-response curve for the FGF21 mutant obtained from quantitative analysis of the data shown in FIG. 17A was shifted to the left compared to the dose-response curve for wild-type FGF21, and the maximum response for the mutant was greater than that for wild-type FGF21 (FIG. 17C). These data show that the Q104M mutant of FGF21 exhibits greater potency and efficacy than native FGF21.

These findings prompted examination of whether the Q104M mutant of FGF21 acts as an FGF21 agonist *in vivo*. Insulin tolerance was used as pharmacodynamic marker, and it was tested whether the mutant can potentiate the hypoglycemic effect of exogenous insulin in mice. As shown in FIGS. 18A and 18C, the Q104M mutant of FGF21 increased the hypoglycemic effect of insulin to a similar degree as wild-type FGF21 did. However, the effect of the mutant persisted for at least twice as long as that of wild-type FGF21, and tended to further increase with time (FIGS. 18A and 18C). These data show that compared to wild-type FGF21, the Q104M mutant of FGF21 has a prolonged potentiating effect on insulin-induced hypoglycemia, which is indicative of agonist potency.

Although the invention has been described in detail for the purposes of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention, which is defined by the following claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 336

<210> SEQ ID NO 1
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

Met Arg Ser Gly Cys Val Val Val His Val Trp Ile Leu Ala Gly Leu
1 5 10 15

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Pro
20 25 30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35 40 45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ala
50 55 60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65 70 75 80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85 90 95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu
100 105 110

Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro
115 120 125

Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu Pro Val Ser
130 135 140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu
145 150 155 160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro
165 170 175

-continued

Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
180 185 190

Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
195 200 205

Val Arg Ser Pro Ser Phe Glu Lys
210 215

<210> SEQ_ID NO 2

<211> LENGTH: 216

<212> TYPE: PRT

<213> ORGANISM: Gorilla gorilla

<400> SEQUENCE: 2

Met Arg Ser Gly Cys Val Val Val His Val Trp Ile Leu Ala Gly Leu
1 5 10 15

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Pro
20 25 30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35 40 45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ala
50 55 60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65 70 75 80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85 90 95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu
100 105 110

Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro
115 120 125

Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu Pro Val Ser
130 135 140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu
145 150 155 160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro
165 170 175

Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
180 185 190

Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
195 200 205

Val Arg Ser Pro Ser Phe Glu Lys
210 215

<210> SEQ_ID NO 3

<211> LENGTH: 216

<212> TYPE: PRT

<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 3

Met Arg Asn Gly Cys Val Val Val His Val Trp Ile Leu Ala Gly Leu
1 5 10 15

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Arg
20 25 30

His Val His Tyr Cys Trp Gly Asp Pro Ile Pro Leu Arg His Leu Tyr
35 40 45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Pro Ala
50 55 60

US 9,475,856 B2

139**140**

-continued

Asn	Cys	Val	Met	Asn	Cys	Ala	Arg	Gly	Gln	Ser	Ala	His	Ser	Leu	Leu
65					70				75						80
Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His															
		85				90				95					
Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu															
		100			105				110						
Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro															
		115			120				125						
Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu Pro Val Ser															
		130			135				140						
Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu															
		145			150				155				160		
Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro															
		165			170				175						
Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu															
		180			185				190						
Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala															
		195			200				205						
Val Arg Ser Pro Ser Phe Glu Lys															
		210			215										

<210> SEQ_ID NO 4

<211> LENGTH: 216

<212> TYPE: PRT

<213> ORGANISM: Macaca mulatta

<400> SEQUENCE: 4

Met	Arg	Ser	Gly	Cys	Val	Val	Val	His	Ala	Trp	Ile	Leu	Ala	Ser	Leu
1					5			10			15				

Trp	Leu	Ala	Val	Ala	Gly	Arg	Pro	Leu	Ala	Phe	Ser	Asp	Ala	Gly	Pro
					20			25			30				

His	Val	His	Tyr	Gly	Trp	Gly	Asp	Pro	Ile	Arg	Leu	Arg	His	Leu	Tyr
					35			40			45				

Thr	Ser	Gly	Pro	His	Gly	Leu	Ser	Ser	Cys	Phe	Leu	Arg	Ile	Arg	Thr
					50			55			60				

Asp	Gly	Val	Val	Asp	Cys	Ala	Arg	Gly	Gln	Ser	Ala	His	Ser	Leu	Leu
					65			70			75				80

Glu	Ile	Lys	Ala	Val	Ala	Leu	Arg	Thr	Val	Ala	Ile	Lys	Gly	Val	His
					85			90			95				

Ser	Val	Arg	Tyr	Leu	Cys	Met	Gly	Ala	Asp	Gly	Lys	Met	Gln	Gly	Leu
					100			105			110				

Leu	Gln	Tyr	Ser	Gl	Gl	Asp	Cys	Ala	Phe	Gl	Gl	Ile	Arg	Pro	
					115			120			125				

Asp	Gly	Tyr	Asn	Val	Tyr	Arg	Ser	Gl	Lys	His	Arg	Leu	Pro	Val	Ser
					130			135			140				

Leu	Ser	Ser	Ala	Lys	Gln	Leu	Tyr	Lys	Asn	Arg	Gly	Leu			
					145			150			155				160

Pro	Leu	Ser	His	Phe	Leu	Pro	Met	Leu	Pro	Met	Ala	Pro	Glu	Glu	Pro
					165			170			175				

Glu	Asp	Leu	Arg	Gly	His	Leu	Gl	U	Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu
					180			185			190					

Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	
					195			200			205					

Val	Arg	Ser	Pro	Ser	Phe	Gl	U	Ser	Asn	Val	Tyr	Arg	Ser	Gl	Lys
					210			215							

-continued

<210> SEQ_ID NO 5
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Pongo abelii

<400> SEQUENCE: 5

```

Met Arg Ser Gly Cys Val Val Val His Ala Trp Ile Leu Ala Gly Leu
1           5          10          15

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ser Gly Pro
20          25          30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35          40          45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ala
50          55          60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65          70          75          80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85          90          95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu
100         105         110

Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro
115         120         125

Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu Pro Val Ser
130         135         140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu
145         150         155         160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro
165         170         175

Glu Asp Leu Arg Arg His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
180         185         190

Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
195         200         205

Val Arg Ser Pro Ser Phe Glu Lys
210         215

```

<210> SEQ_ID NO 6
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Nomascus leucogenys

<400> SEQUENCE: 6

```

Met Arg Ser Glu Cys Val Val Val His Ala Trp Ile Leu Ala Gly Leu
1           5          10          15

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Pro
20          25          30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35          40          45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ala
50          55          60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65          70          75          80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85          90          95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu
100         105         110

```

-continued

Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Glu Ile Arg Pro
 115 120 125
 Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu Pro Val Ser
 130 135 140
 Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu
 145 150 155 160
 Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro
 165 170 175
 Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
 180 185 190
 Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
 195 200 205
 Val Arg Ser Pro Ser Phe Glu Lys
 210 215

<210> SEQ_ID NO 7
 <211> LENGTH: 142
 <212> TYPE: PRT
 <213> ORGANISM: Callithrix jacchus

<400> SEQUENCE: 7

Met Trp Lys Ala Thr Ala Gly Gly Gln Gln Gly Gln Ser Glu Ala Gln
 1 5 10 15
 Met Ser Thr Cys Pro His Val Pro Arg Pro Leu Trp Ile Ala Gln Ser
 20 25 30
 Cys Leu Phe Ser Leu Gln Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe
 35 40 45
 Glu Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr Trp Ser Glu Lys
 50 55 60
 His Arg Leu Pro Val Ser Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr
 65 70 75 80
 Lys Lys Arg Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro
 85 90 95
 Ile Ala Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp
 100 105 110
 Val Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu
 115 120 125
 Val Thr Gly Leu Glu Ala Val Asn Ser Pro Ser Phe Glu Lys
 130 135 140

<210> SEQ_ID NO 8
 <211> LENGTH: 219
 <212> TYPE: PRT
 <213> ORGANISM: Microcebus murinus

<400> SEQUENCE: 8

Met Pro Ser Gly Gln Ser Gly Val Ala Ala Arg Ala Leu Ile Leu
 1 5 10 15
 Ala Gly Leu Trp Leu Thr Ala Ala Gly Arg Pro Leu Ala Phe Ser Asp
 20 25 30
 Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg
 35 40 45
 His Leu Tyr Thr Ala Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg
 50 55 60
 Ile Arg Ala Asp Gly Ser Val Asp Cys Ala Arg Gly Gln Ser Ala His
 65 70 75 80

-continued

Ser Leu Leu Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala Ile Lys
85 90 95

Gly Val His Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met
100 105 110

Gln Gly Leu Leu Arg Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Glu
115 120 125

Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu
130 135 140

Pro Val Ser Leu Ser Ser Ala Arg Gln Arg Gln Leu Tyr Lys Gly Arg
145 150 155 160

Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Val Thr Pro
165 170 175

Ala Glu Thr Gly Asp Leu Arg Asp His Leu Glu Ser Asp Met Phe Ala
180 185 190

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Thr Arg
195 200 205

Leu Gly Val Val Lys Ser Pro Ser Phe Gln Lys
210 215

<210> SEQ_ID NO 9

<211> LENGTH: 138

<212> TYPE: PRT

<213> ORGANISM: Choloepus hoffmanni

<400> SEQUENCE: 9

Leu Leu Glu Met Lys Ala Val Ala Leu Arg Ala Val Ala Ile Lys Gly
1 5 10 15

Val His Ser Ala Leu Tyr Leu Cys Met Asn Ala Asp Gly Ser Leu His
20 25 30

Gly Leu Pro Arg Tyr Ser Ala Glu Asp Cys Ala Phe Glu Glu Glu Ile
35 40 45

Arg Pro Asp Gly Tyr Asn Val Tyr Trp Ser Arg Lys His Gly Leu Pro
50 55 60

Val Ser Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Gly Arg Gly
65 70 75 80

Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Thr Pro Ala
85 90 95

Glu Pro Ala Asp Pro Gly Asp Asp Val Glu Ser Asp Met Phe Ser Ser
100 105 110

Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Ser Arg Leu
115 120 125

Glu Leu Val Asn Ser Pro Ser Phe Gln Thr
130 135

<210> SEQ_ID NO 10

<211> LENGTH: 205

<212> TYPE: PRT

<213> ORGANISM: Ailuropoda melanoleuca

<400> SEQUENCE: 10

Val Leu Ala Gly Leu Cys Leu Ala Val Ala Gly Arg Pro Leu Ala Phe
1 5 10 15

Ser Asp Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Pro Ile Arg
20 25 30

Leu Arg His Leu Tyr Thr Ala Gly Pro His Gly Leu Ser Ser Cys Phe
35 40 45

-continued

Leu Arg Ile Arg Ala Asp Gly Gly Val Asp Cys Ala Arg Gly Gln Ser
 50 55 60
 Ala His Ser Leu Val Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala
 65 70 75 80
 Ile Lys Gly Val His Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly
 85 90 95
 Arg Met Gln Gly Leu Pro Gln Tyr Ser Ala Gly Asp Cys Ala Phe Glu
 100 105 110
 Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg Ser Lys Lys His
 115 120 125
 Arg Leu Pro Val Ser Leu Ser Gly Ala Lys Gln Arg Gln Leu Tyr Lys
 130 135 140
 Asp Arg Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Gly
 145 150 155 160
 Ser Pro Ala Glu Pro Arg Asp Leu Gln Asp His Ala Glu Ser Asp Gly
 165 170 175
 Phe Ser Ala Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala
 180 185 190
 Thr Lys Met Gly Leu Val Lys Ser Pro Ser Phe Gln Lys
 195 200 205

<210> SEQ ID NO 11
 <211> LENGTH: 218
 <212> TYPE: PRT
 <213> ORGANISM: Sus scrofa
 <400> SEQUENCE: 11

Met Arg Ser Ala Pro Ser Arg Cys Ala Val Val Arg Ala Leu Val Leu
 1 5 10 15
 Ala Gly Leu Trp Leu Ala Ala Ala Gly Arg Pro Leu Ala Phe Ser Asp
 20 25 30
 Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Ser Val Arg Leu Arg
 35 40 45
 His Leu Tyr Thr Ala Ser Pro His Gly Val Ser Ser Cys Phe Leu Arg
 50 55 60
 Ile His Ser Asp Gly Pro Val Asp Cys Ala Pro Gly Gln Ser Ala His
 65 70 75 80
 Ser Leu Met Glu Ile Arg Ala Val Ala Leu Ser Thr Val Ala Ile Lys
 85 90 95
 Gly Glu Arg Ser Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln
 100 105 110
 Gly Gln Thr Gln Tyr Ser Asp Glu Asp Cys Ala Phe Glu Glu Ile
 115 120 125
 Arg Pro Asp Gly Tyr Asn Val Tyr Trp Ser Lys Lys His His Leu Pro
 130 135 140
 Val Ser Leu Ser Ser Ala Arg Gln Arg Gln Leu Tyr Lys Gly Arg Gly
 145 150 155 160
 Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Ser Thr Leu Pro Ala
 165 170 175
 Glu Pro Glu Asp Leu Gln Asp Pro Phe Lys Ser Asp Leu Phe Ser Leu
 180 185 190
 Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Arg Ile Ala Ala Lys Leu
 195 200 205
 Gly Ala Val Lys Ser Pro Ser Phe Tyr Lys

US 9,475,856 B2

149

-continued

210

215

<210> SEQ ID NO 12

<211> LENGTH: 218

<212> TYPE: PRT

<213> ORGANISM: Bos Taurus

<400> SEQUENCE: 12

Met	Arg	Ser	Ala	Pro	Ser	Arg	Cys	Ala	Val	Ala	Arg	Ala	Leu	Val	Leu
1									5	10			15		

Ala	Gly	Leu	Trp	Leu	Ala	Ala	Ala	Gly	Arg	Pro	Leu	Ala	Phe	Ser	Asp
									20	25			30		

Ala	Gly	Pro	His	Val	His	Tyr	Gly	Trp	Gly	Glu	Ser	Val	Arg	Leu	Arg
									35	40			45		

His	Leu	Tyr	Thr	Ala	Gly	Pro	Gln	Gly	Leu	Tyr	Ser	Cys	Phe	Leu	Arg
									50	55			60		

Ile	His	Ser	Asp	Gly	Ala	Val	Asp	Cys	Ala	Gln	Val	Gln	Ser	Ala	His
65									70	75			80		

Ser	Leu	Met	Glu	Ile	Arg	Ala	Val	Ala	Leu	Ser	Thr	Val	Ala	Ile	Lys
									85	90			95		

Gly	Glu	Arg	Ser	Val	Leu	Tyr	Leu	Cys	Met	Asp	Ala	Asp	Gly	Lys	Met
									100	105			110		

Gln	Gly	Leu	Thr	Gln	Tyr	Ser	Ala	Glu	Asp	Cys	Ala	Phe	Glu	Glu	
									115	120			125		

Ile	Arg	Pro	Asp	Gly	Tyr	Asn	Val	Tyr	Trp	Ser	Arg	Lys	His	His	Leu
130									135	140					

Pro	Val	Ser	Leu	Ser	Ser	Ser	Arg	Gln	Arg	Gln	Leu	Phe	Lys	Ser	Arg
145								150		155			160		

Gly	Phe	Leu	Pro	Leu	Ser	His	Phe	Leu	Pro	Met	Leu	Ser	Thr	Ile	Pro
									165	170			175		

Ala	Glu	Pro	Glu	Asp	Leu	Gln	Glu	Pro	Leu	Lys	Pro	Asp	Phe	Phe	Leu
									180	185			190		

Pro	Leu	Lys	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Ala	Thr	Lys	Leu
								195	200			205			

Gly	Ser	Val	Lys	Ser	Pro	Ser	Phe	Tyr	Asn						
									210	215					

<210> SEQ ID NO 13

<211> LENGTH: 192

<212> TYPE: PRT

<213> ORGANISM: Canis lupus familiaris

<400> SEQUENCE: 13

Leu	Ala	Phe	Ser	Asp	Ala	Gly	Pro	His	Val	His	Ser	Phe	Trp	Gly	Glu
1									5	10			15		

Pro	Ile	Arg	Leu	Arg	His	Leu	Tyr	Thr	Ala	Gly	Pro	His	Gly	Leu	Ser
								20	25			30			

Ser	Cys	Phe	Leu	Arg	Ile	Arg	Ala	Asp	Gly	Gly	Val	Asp	Cys	Ala	Arg
									35	40			45		

Gly	Gln	Ser	Ala	His	Ser	Leu	Met	Glu	Met	Arg	Ala	Val	Ala	Leu	Arg
									50	55			60		

Thr	Val	Ala	Ile	Lys	Gly	Val	His	Ser	Gly	Arg	Tyr	Leu	Cys	Met	Gly
65									70	75			80		

Ala	Asp	Gly	Arg	Met	Gln	Gly	Leu	Pro	Gln	Tyr	Ser	Ala	Gly	Asp	Cys
									85	90			95		

Thr	Phe	Glu	Glu	Ile	Arg	Pro	Asp	Gly	Tyr	Asn	Val	Tyr	Trp	Ser	
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

149**150**

US 9,475,856 B2

151**152**

-continued

100	105	110
Lys Lys His His Leu Pro Ile Ser Leu Ser Ser Ala Lys Gln Arg Gln		
115	120	125
Leu Tyr Lys Gly Arg Gly Phe Leu Pro Leu Ser His Phe Leu Pro Ile		
130	135	140
Leu Pro Gly Ser Pro Thr Glu Pro Arg Asp Leu Glu Asp His Val Glu		
145	150	155
Ser Asp Gly Phe Ser Ala Ser Leu Glu Thr Asp Ser Met Asp Pro Phe		
165	170	175
Gly Ile Ala Thr Lys Ile Gly Leu Val Lys Ser Pro Ser Phe Gln Lys		
180	185	190

<210> SEQ ID NO 14
<211> LENGTH: 218
<212> TYPE: PRT
<213> ORGANISM: Oryctolagus cuniculus

<400> SEQUENCE: 14

Met Arg Arg Ala Pro Ser Gly Gly Ala Ala Ala Arg Ala Leu Val Leu		
1	5	10
Ala Gly Leu Trp Leu Ala Ala Ala Arg Pro Leu Ala Leu Ser Asp		
20	25	30
Ala Gly Pro His Leu His Tyr Gly Trp Gly Glu Pro Val Arg Leu Arg		
35	40	45
His Leu Tyr Ala Thr Ser Ala His Gly Val Ser His Cys Phe Leu Arg		
50	55	60
Ile Arg Ala Asp Gly Ala Val Asp Cys Glu Arg Ser Gln Ser Ala His		
65	70	75
Ser Leu Leu Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala Phe Lys		
85	90	95
Gly Val His Ser Ser Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met		
100	105	110
Arg Gly Gln Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Gln Glu Glu		
115	120	125
Ile Ser Ser Gly Tyr Asn Val Tyr Arg Ser Thr Thr His His Leu Pro		
130	135	140
Val Ser Leu Ser Ser Ala Lys Gln Arg His Leu Tyr Lys Thr Arg Gly		
145	150	155
Phe Leu Pro Leu Ser His Phe Leu Pro Val Leu Pro Leu Ala Ser Glu		
165	170	175
Glu Thr Ala Ala Leu Gly Asp His Pro Glu Ala Asp Leu Phe Ser Pro		
180	185	190
Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Met Ala Thr Lys Leu		
195	200	205
Gly Pro Val Lys Ser Pro Ser Phe Gln Lys		
210	215	

<210> SEQ ID NO 15
<211> LENGTH: 215
<212> TYPE: PRT
<213> ORGANISM: Pteropus vampyrus

<400> SEQUENCE: 15

Met Arg Ser Pro Cys Ala Val Ala Arg Ala Leu Val Ala Gly Leu		
1	5	10
Trp Leu Ala Ser Ala Ala Gly Pro Leu Ala Leu Ser Asp Ala Gly Pro		

US 9,475,856 B2

153

-continued

154

20	25	30
His Val His Tyr Gly Trp Gly Glu Ala Ile Arg Leu Arg His Leu Tyr		
35	40	45
Thr Ala Gly Pro His Gly Pro Ser Ser Cys Phe Leu Arg Ile Arg Ala		
50	55	60
Asp Gly Ala Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Val		
65	70	75
Glu Ile Arg Ala Val Ala Leu Arg Asn Val Ala Ile Lys Gly Val His		
85	90	95
Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met Leu Gly Leu		
100	105	110
Leu Gln Tyr Ser Ala Asp Asp Cys Ala Phe Glu Glu Ile Arg Pro		
115	120	125
Asp Gly Tyr Asn Val Tyr His Ser Lys Lys His His Leu Pro Val Ser		
130	135	140
Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asp Arg Gly Phe Leu		
145	150	155
Pro Leu Ser His Phe Leu Pro Met Leu Pro Arg Ser Pro Thr Glu Pro		
165	170	175
Glu Asn Phe Glu Asp His Leu Glu Ala Asp Thr Phe Ser Ser Leu Glu		
180	185	190
Thr Asp Asp Met Asp Pro Phe Gly Ile Ala Ser Lys Leu Gly Leu Glu		
195	200	205
Glu Ser Pro Ser Phe Gln Lys		
210	215	

<210> SEQ_ID NO 16

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: Tursiops truncates

<400> SEQUENCE: 16

Met Arg Ser Ala Pro Ser Arg Cys Ala Val Ala Arg Ala Leu Val Leu		
1	5	10
Ala Gly Leu Trp Leu Ala Ala Ala Gly Arg Pro Leu Ala Phe Ser Asp		
20	25	30
Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Ser Val Arg Leu Arg		
35	40	45
His Leu Tyr Thr Ala Gly Pro Gln Gly Leu Ser Ser Cys Phe Leu Arg		
50	55	60
Ile His Ser Asp Gly Ala Val Asp Cys Ala Pro Val Gln Ser Ala His		
65	70	75
Ser Leu Met Glu Ile Arg Ala Val Ala Leu Ser Thr Val Ala Ile Lys		
85	90	95
Gly Glu Arg Ser Val Leu Tyr Leu Cys Met Gly Ala Asp Gly Lys Met		
100	105	110
Gln Gly Leu Ser Gln Tyr Ser Ala Glu Asp Cys Ala Phe Glu Glu		
115	120	125
Ile Arg Pro Asp Gly Tyr Asn Val Tyr Trp Ser Lys Lys His His Leu		
130	135	140
Pro Val Ser Leu Ser Ser Ala Arg Gln Arg Gln Leu Phe Lys Gly Arg		
145	150	155
Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Ser Thr Ile Pro		
165	170	175

US 9,475,856 B2

155

-continued

Thr Glu Pro Asp Glu Ile Gln Asp His Leu Lys Pro Asp Leu Phe Ala
 180 185 190

Leu Pro Leu Lys Thr Asp Ser Met Asp Pro Phe Gly Leu Ala Thr Lys
 195 200 205

Leu Gly Val Val Lys Ser Pro Ser Phe Tyr Lys
 210 215

<210> SEQ ID NO 17

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: Myotis lucifugus

<400> SEQUENCE: 17

Met Gln Ser Ala Trp Ser Arg Arg Val Val Ala Arg Ala Leu Val Leu
 1 5 10 15

Ala Ser Leu Gly Leu Ala Ser Ala Gly Gly Pro Leu Gly Leu Ser Asp
 20 25 30

Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Ser Ile Arg Leu Arg
 35 40 45

His Leu Tyr Thr Ser Gly Pro His Gly Pro Ser Ser Cys Phe Leu Arg
 50 55 60

Ile Arg Ala Asp Gly Ala Val Asp Cys Ala Arg Gly Gln Ser Ala His
 65 70 75 80

Ser Leu Val Glu Ile Arg Ala Val Ala Leu Arg Lys Val Ala Ile Lys
 85 90 95

Gly Val His Ser Ala Leu Tyr Leu Cys Met Gly Gly Asp Gly Arg Met
 100 105 110

Leu Gly Leu Pro Gln Phe Ser Pro Glu Asp Cys Ala Phe Glu Glu Glu
 115 120 125

Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg Ser Gln Lys His Gln Leu
 130 135 140

Pro Val Ser Leu Ser Ser Ala Arg Gln Arg Gln Leu Phe Lys Ala Arg
 145 150 155 160

Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Ser Ser Pro
 165 170 175

Ala Gly Pro Val Pro Arg Glu Arg Pro Ser Glu Pro Asp Glu Phe Ser
 180 185 190

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Asn Asn
 195 200 205

Leu Arg Leu Val Arg Ser Pro Ser Phe Gln Glu
 210 215

<210> SEQ ID NO 18

<211> LENGTH: 185

<212> TYPE: PRT

<213> ORGANISM: Ornithorhynchus anatinus

<400> SEQUENCE: 18

Met Leu Ser Cys Val Val Leu Pro Ser Leu Leu Glu Ile Lys Ala Val
 1 5 10 15

Ala Val Arg Thr Val Ala Ile Lys Gly Val His Ile Ser Arg Tyr Leu
 20 25 30

Cys Met Glu Glu Asp Gly Lys Thr Pro Trp Ala Arg Leu Leu Glu Ile
 35 40 45

Lys Ala Val Ala Val Arg Thr Val Ala Ile Lys Gly Val His Ser Ser
 50 55 60

156

US 9,475,856 B2

157**158**

-continued

Arg	Tyr	Leu	Cys	Met	Glu	Glu	Asp	Gly	Lys	Leu	His	Gly	Gln	Ile	Trp
65				70					75					80	
Tyr	Ser	Ala	Glu	Asp	Cys	Ala	Phe	Glu	Glu	Glu	Ile	Arg	Pro	Asp	Gly
					85				90					95	
Tyr	Asn	Val	Tyr	Lys	Ser	Lys	Lys	Tyr	Gly	Val	Pro	Val	Ser	Leu	Ser
					100				105					110	
Ser	Ala	Lys	Gln	Arg	Gln	Gln	Phe	Lys	Gly	Arg	Asp	Phe	Leu	Pro	Leu
					115			120					125		
Ser	Arg	Phe	Leu	Pro	Met	Ile	Asn	Thr	Val	Pro	Val	Glu	Pro	Ala	Glu
					130			135				140			
Phe	Gly	Asp	Tyr	Ala	Asp	Tyr	Phe	Glu	Ser	Asp	Ile	Phe	Ser	Ser	Pro
					145			150			155			160	
Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Arg	Ile	Ala	Pro	Lys	Leu	Ser
					165			170					175		
Pro	Val	Lys	Ser	Pro	Ser	Phe	Gln	Lys							
					180			185							

<210> SEQ ID NO 19

<211> LENGTH: 212

<212> TYPE: PRT

<213> ORGANISM: Monodelphis domestica

<400> SEQUENCE: 19

Met	Ala	Gln	Leu	Leu	Ala	Pro	Leu	Leu	Thr	Leu	Ala	Ala	Leu	Trp	Leu
1					5				10					15	
Ala	Pro	Thr	Ala	Arg	Ala	Arg	Pro	Leu	Val	Asp	Ala	Gly	Pro	His	Val
					20				25					30	
Tyr	Tyr	Gly	Trp	Gly	Glu	Pro	Ile	Arg	Leu	Arg	His	Leu	Tyr	Thr	Ala
					35			40			45				
Asn	Arg	His	Gly	Leu	Ala	Ser	Phe	Ser	Phe	Leu	Arg	Ile	His	Arg	Asp
					50			55			60				
Gly	Arg	Val	Asp	Gly	Ser	Arg	Ser	Gln	Ser	Ala	Leu	Ser	Leu	Leu	Glu
					65			70			75			80	
Ile	Lys	Ala	Val	Ala	Leu	Arg	Met	Val	Ala	Ile	Lys	Gly	Val	His	Ser
					85			90			95				
Ser	Arg	Tyr	Leu	Cys	Met	Gly	Asp	Ala	Gly	Lys	Leu	Gln	Gly	Ser	Val
					100			105			110				
Arg	Phe	Ser	Ala	Glu	Asp	Cys	Thr	Phe	Glu	Glu	Gln	Ile	Arg	Pro	Asp
					115			120			125				
Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Pro	Lys	Tyr	Asn	Leu	Pro	Val	Ser	Leu
					130			135			140				
Cys	Thr	Asp	Lys	Gln	Arg	Gln	Gln	Ala	His	Gly	Lys	Glu	His	Leu	Pro
					145			150			155			160	
Leu	Ser	His	Phe	Leu	Pro	Met	Ile	Asn	Ala	Ile	Pro	Leu	Ala	Glu	
					165			170			175				
Glu	Pro	Glu	Gly	Pro	Arg	Met	Leu	Ala	Ala	Pro	Leu	Glu	Thr	Asp	Ser
					180			185			190				
Met	Asp	Pro	Phe	Gly	Leu	Thr	Ser	Lys	Leu	Leu	Pro	Val	Lys	Ser	Pro
					195			200			205				
Ser	Phe	Gln	Lys												
					210										

<210> SEQ ID NO 20

<211> LENGTH: 220

<212> TYPE: PRT

<213> ORGANISM: Anolis carolinensis

US 9,475,856 B2

159

-continued

<400> SEQUENCE: 20

```

Met Cys Arg Arg Ala Leu Pro Leu Leu Gly Ala Leu Leu Gly Leu Ala
1           5          10          15

Ala Val Ala Ser Arg Ala Leu Pro Leu Thr Asp Ala Gly Pro His Val
20          25          30

Ser Tyr Gly Trp Gly Glu Pro Val Arg Leu Arg His Leu Tyr Thr Ala
35          40          45

Gly Arg Gln Gly Leu Phe Ser Gln Phe Leu Arg Ile His Ala Asp Gly
50          55          60

Arg Val Asp Gly Ala Gly Ser Gln Asn Arg Gln Ser Leu Leu Glu Ile
65          70          75          80

Arg Ala Val Ser Leu Arg Ala Val Ala Leu Lys Gly Val His Ser Ser
85          90          95

Arg Tyr Leu Cys Met Glu Glu Asp Gly Arg Leu Arg Gly Met Leu Arg
100         105         110

Tyr Ser Ala Glu Asp Cys Ser Phe Glu Glu Glu Met Arg Pro Asp Gly
115         120         125

Tyr Asn Ile Tyr Lys Ser Lys Lys Tyr Gly Val Leu Val Ser Leu Ser
130         135         140

Asn Ala Arg Gln Arg Gln Phe Lys Gly Lys Asp Phe Leu Pro Leu
145         150         155         160

Ser His Phe Leu Pro Met Ile Asn Thr Val Pro Val Glu Ser Ala Asp
165         170         175

Phe Gly Glu Tyr Gly Asp Thr Arg Gln His Tyr Glu Ser Asp Ile Phe
180         185         190

Ser Ser Arg Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Thr Ser
195         200         205

Glu Val Ser Ser Val Gln Ser Pro Ser Phe Gly Lys
210         215         220

```

<210> SEQ ID NO 21

<211> LENGTH: 179

<212> TYPE: PRT

<213> ORGANISM: Ochotona princeps

<400> SEQUENCE: 21

```

Val Arg Ser Arg Gly Ala Met Ala Arg Ala Leu Val Leu Ala Thr Leu
1           5          10          15

Trp Leu Ala Ala Thr Gly Arg Pro Leu Ala Leu Ser Asp Ala Gly Pro
20          25          30

His Leu His Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg His Leu Tyr
35          40          45

Ala Thr Ser Ala His Gly Leu Ser His Cys Phe Leu Arg Ile Arg Thr
50          55          60

Asp Gly Thr Val Asp Cys Glu Arg Ser Gln Ser Ala His Leu Gln Tyr
65          70          75          80

Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Ser Ser Gly Tyr Asn
85          90          95

Val Tyr Arg Ser Arg Arg Tyr Gln Leu Pro Val Ser Leu Gly Ser Ala
100         105         110

Arg Gln Arg Gln Leu Gln Arg Ser Arg Gly Phe Leu Pro Leu Ser His
115         120         125

Phe Leu Pro Val Leu Pro Ala Ala Ser Glu Glu Val Ala Ala Pro Ala
130         135         140

```

160

-continued

Asp His Pro Gln Ala Asp Pro Phe Ser Pro Leu Glu Thr Asp Ser Met
 145 150 155 160

Asp Pro Phe Gly Met Ala Thr Lys Arg Gly Leu Val Lys Ser Pro Ser
 165 170 175

Phe Gln Lys

<210> SEQ ID NO 22

<211> LENGTH: 221

<212> TYPE: PRT

<213> ORGANISM: Cavia porcellus

<400> SEQUENCE: 22

Met Trp Ser Ala Pro Ser Gly Cys Val Val Ile Arg Ala Leu Val Leu
 1 5 10 15

Ala Gly Leu Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Arg Arg Ser
 20 25 30

Leu Ala Leu Ser Asp Gln Gly Pro His Leu Tyr Tyr Gly Trp Asp Gln
 35 40 45

Pro Ile Arg Leu Arg His Leu Tyr Ala Ala Gly Pro Tyr Gly Arg Ser
 50 55 60

Arg Cys Phe Leu Arg Ile His Thr Asp Gly Ala Val Asp Cys Val Glu
 65 70 75 80

Glu Gln Ser Glu His Cys Leu Leu Glu Ile Arg Ala Val Ala Leu Glu
 85 90 95

Thr Val Ala Ile Lys Asp Ile Asn Ser Val Arg Tyr Leu Cys Met Gly
 100 105 110

Pro Asp Gly Arg Met Arg Gly Leu Pro Trp Tyr Ser Glu Glu Asp Cys
 115 120 125

Ala Phe Lys Glu Glu Ile Ser Tyr Pro Gly Tyr Ser Val Tyr Arg Ser
 130 135 140

Gln Lys His His Leu Pro Ile Val Leu Ser Ser Val Lys Gln Arg Gln
 145 150 155 160

Gln Tyr Gln Ser Lys Gly Val Val Pro Leu Ser Tyr Phe Leu Pro Met
 165 170 175

Leu Pro Lys Ala Ser Val Glu Pro Ser Asp Glu Glu Ser Ser Val
 180 185 190

Phe Ser Leu Pro Leu Lys Thr Asp Ser Met Asp Pro Phe Gly Met Ala
 195 200 205

Ser Glu Ile Gly Leu Val Lys Ser Pro Ser Phe Gln Lys
 210 215 220

<210> SEQ ID NO 23

<211> LENGTH: 196

<212> TYPE: PRT

<213> ORGANISM: Tupaia belangeri

<400> SEQUENCE: 23

Met Arg Arg Thr Pro Ser Gly Phe Ala Val Ala Arg Val Leu Phe Leu
 1 5 10 15

Gly Ser Leu Trp Leu Ala Ala Gly Ser Pro Leu Ala Leu Ser Asp
 20 25 30

Ala Gly Pro His Val Asn Tyr Gly Trp Asp Glu Ser Ile Arg Leu Arg
 35 40 45

His Leu Tyr Thr Ala Ser Pro His Gly Ser Thr Ser Cys Phe Leu Arg
 50 55 60

US 9,475,856 B2

163**164**

-continued

Ile	Arg	Asp	Asp	Gly	Ser	Val	Asp	Cys	Ala	Arg	Gly	Gln	Ser	Leu	His
65				70				75						80	
Ser	Leu	Leu	Glu	Ile	Lys	Ala	Val	Ala	Leu	Gln	Thr	Val	Ala	Ile	Lys
			85				90					95			
Gly	Val	Tyr	Ser	Val	Arg	Tyr	Leu	Cys	Met	Asp	Ala	Asp	Gly	Arg	Met
			100				105					110			
Gln	Gly	Leu	Ser	Thr	Lys	His	Gly	Leu	Pro	Val	Ser	Leu	Ser	Ser	Ala
	115				120						125				
Lys	Gln	Arg	Gln	Leu	Leu	Thr	Val	Arg	Gly	Phe	Pro	Ser	Leu	Pro	His
	130				135			140							
Phe	Leu	Leu	Met	Met	Ala	Lys	Thr	Ser	Ala	Gly	Pro	Gly	Asn	Pro	Arg
145					150			155					160		
Asp	His	Pro	Gly	Ser	Asn	Thr	Phe	Ser	Leu	Pro	Leu	Glu	Thr	Asp	Ser
	165					170						175			
Met	Asp	Pro	Phe	Gly	Met	Thr	Thr	Arg	His	Gly	Leu	Val	Lys	Ser	Pro
	180					185			190						
Ser	Phe	Gln	Asn												
	195														

<210> SEQ ID NO 24

<211> LENGTH: 218

<212> TYPE: PRT

<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 24

Met	Ala	Arg	Lys	Trp	Ser	Gly	Arg	Ile	Val	Ala	Arg	Ala	Leu	Val	Leu
1				5				10					15		
Ala	Thr	Leu	Trp	Leu	Ala	Val	Ser	Gly	Arg	Pro	Leu	Val	Gln	Gln	Ser
	20				25				30						
Gln	Ser	Val	Ser	Asp	Glu	Gly	Pro	Leu	Phe	Leu	Tyr	Gly	Trp	Gly	Lys
	35				40				45						
Ile	Thr	Arg	Leu	Gln	Tyr	Leu	Tyr	Ser	Ala	Gly	Pro	Tyr	Val	Ser	Asn
	50				55			60							
Cys	Phe	Leu	Arg	Ile	Arg	Ser	Asp	Gly	Ser	Val	Asp	Cys	Glu	Glu	Asp
65					70			75					80		
Gln	Asn	Glu	Arg	Asn	Leu	Leu	Glu	Phe	Arg	Ala	Val	Ala	Leu	Lys	Thr
	85					90			95						
Ile	Ala	Ile	Lys	Asp	Val	Ser	Ser	Val	Arg	Tyr	Leu	Cys	Met	Ser	Ala
	100				105			110							
Asp	Gly	Lys	Ile	Tyr	Gly	Leu	Ile	Arg	Tyr	Ser	Glu	Glu	Asp	Cys	Thr
	115				120			125							
Phe	Arg	Glu	Glu	Met	Asp	Cys	Leu	Gly	Tyr	Asn	Gln	Tyr	Arg	Ser	Met
130					135			140							
Lys	His	His	Leu	His	Ile	Ile	Phe	Ile	Lys	Ala	Lys	Pro	Arg	Glu	Gln
145					150			155				160			
Leu	Gln	Gly	Gln	Lys	Pro	Ser	Asn	Phe	Ile	Pro	Ile	Phe	His	Arg	Ser
	165					170			175						
Phe	Phe	Glu	Ser	Thr	Asp	Gln	Leu	Arg	Ser	Lys	Met	Phe	Ser	Leu	Pro
	180					185			190						
Leu	Glu	Ser	Asp	Ser	Met	Asp	Pro	Phe	Arg	Met	Val	Glu	Asp	Val	Asp
	195					200			205						
His	Leu	Val	Lys	Ser	Pro	Ser	Phe	Gln	Lys						
	210				215										

<210> SEQ ID NO 25

US 9,475,856 B2

165

-continued

<211> LENGTH: 218
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 25

```

Met Ala Arg Lys Trp Asn Gly Arg Ala Val Ala Arg Ala Leu Val Leu
1           5          10          15

Ala Thr Leu Trp Leu Ala Val Ser Gly Arg Pro Leu Ala Gln Gln Ser
20          25          30

Gln Ser Val Ser Asp Glu Asp Pro Leu Phe Leu Tyr Gly Trp Gly Lys
35          40          45

Ile Thr Arg Leu Gln Tyr Leu Tyr Ser Ala Gly Pro Tyr Val Ser Asn
50          55          60

Cys Phe Leu Arg Ile Arg Ser Asp Gly Ser Val Asp Cys Glu Glu Asp
65          70          75          80

Gln Asn Glu Arg Asn Leu Leu Glu Phe Arg Ala Val Ala Leu Lys Thr
85          90          95

Ile Ala Ile Lys Asp Val Ser Ser Val Arg Tyr Leu Cys Met Ser Ala
100         105         110

Asp Gly Lys Ile Tyr Gly Leu Ile Arg Tyr Ser Glu Glu Asp Cys Thr
115         120         125

Phe Arg Glu Glu Met Asp Cys Leu Gly Tyr Asn Gln Tyr Arg Ser Met
130         135         140

Lys His His Leu His Ile Ile Phe Ile Gln Ala Lys Pro Arg Glu Gln
145         150         155         160

Leu Gln Asp Gln Lys Pro Ser Asn Phe Ile Pro Val Phe His Arg Ser
165         170         175

Phe Phe Glu Thr Gly Asp Gln Leu Arg Ser Lys Met Phe Ser Leu Pro
180         185         190

Leu Glu Ser Asp Ser Met Asp Pro Phe Arg Met Val Glu Asp Val Asp
195         200         205

His Leu Val Lys Ser Pro Ser Phe Gln Lys
210         215

```

<210> SEQ ID NO 26
<211> LENGTH: 227
<212> TYPE: PRT
<213> ORGANISM: Gallus gallus

<400> SEQUENCE: 26

```

Met Gly Pro Ala Arg Pro Ala Ala Pro Gly Ala Ala Leu Ala Leu Leu
1           5          10          15

Gly Ile Ala Ala Ala Ala Ala Ala Arg Ser Leu Pro Leu Pro Asp
20          25          30

Val Gly Gly Pro His Val Asn Tyr Gly Trp Gly Glu Pro Ile Arg Leu
35          40          45

Arg His Leu Leu His Arg Pro Gly Lys His Gly Leu Phe Ser Cys Phe
50          55          60

Leu Arg Ile Gly Gly Asp Gly Arg Val Asp Ala Val Gly Ser Gln Ser
65          70          75          80

Pro Gln Ser Leu Leu Glu Ile Arg Ala Val Ala Val Arg Thr Val Ala
85          90          95

Ile Lys Gly Val Gln Ser Ser Arg Tyr Leu Cys Met Asp Glu Ala Gly
100         105         110

Arg Leu His Gly Gln Leu Ser Tyr Ser Ile Glu Asp Cys Ser Phe Glu
115         120         125

```

166

-continued

Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr Lys Ser Lys Lys Tyr
 130 135 140

Gly Ile Ser Val Ser Leu Ser Ser Ala Lys Gln Arg Gln Gln Phe Lys
 145 150 155 160

Gly Lys Asp Phe Leu Pro Leu Ser His Phe Leu Pro Met Ile Asn Thr
 165 170 175

Val Pro Val Glu Val Thr Asp Phe Gly Glu Tyr Gly Asp Tyr Ser Gln
 180 185 190

Ala Phe Glu Pro Glu Val Tyr Ser Ser Pro Leu Glu Thr Asp Ser Met
 195 200 205

Asp Pro Phe Gly Ile Thr Ser Lys Leu Ser Pro Val Lys Ser Pro Ser
 210 215 220

Phe Gln Lys
 225

<210> SEQ_ID NO 27
 <211> LENGTH: 237
 <212> TYPE: PRT
 <213> ORGANISM: Taeniopygia guttata

<400> SEQUENCE: 27

Met Val Ile Ile Ser Asn Leu Tyr Leu Met Gln Asn Asp Val Met Met
 1 5 10 15

Asn Met Arg Arg Ala Pro Leu Arg Val His Ala Ala Arg Ser Ser Ala
 20 25 30

Thr Pro Ala Ser Ala Leu Pro Leu Pro Pro Asp Ala Gly Pro His
 35 40 45

Leu Lys Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg His Leu Tyr Thr
 50 55 60

Ala Ser Lys His Gly Leu Phe Ser Cys Phe Leu Arg Ile Gly Ala Asp
 65 70 75 80

Gly Arg Val Asp Ala Ala Gly Ser Gln Ser Pro Gln Ser Leu Leu Glu
 85 90 95

Ile Arg Ala Val Ala Val Arg Thr Val Ala Ile Lys Gly Val Gln Ser
 100 105 110

Ser Arg Tyr Leu Cys Met Asp Glu Ala Gly Arg Leu His Gly Gln Leu
 115 120 125

Arg Asn Ser Thr Glu Asp Cys Ser Phe Glu Glu Ile Arg Pro Asp
 130 135 140

Gly Tyr Asn Val Tyr Arg Ser Lys Lys His Gly Ile Ser Val Ser Leu
 145 150 155 160

Ser Ser Ala Lys Gln Arg Gln Gln Phe Lys Gly Lys Asp Phe Leu Pro
 165 170 175

Leu Ser His Phe Leu Pro Met Ile Asn Thr Val Pro Met Glu Ser Ala
 180 185 190

Asp Phe Gly Glu Tyr Gly Asp Tyr Ser Gln Ala Phe Glu Ala Glu Ala
 195 200 205

Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala
 210 215 220

Ser Lys Leu Ser Leu Val Lys Ser Pro Ser Phe Gln Asn
 225 230 235

<210> SEQ_ID NO 28
 <211> LENGTH: 210
 <212> TYPE: PRT

US 9,475,856 B2

169

-continued

<213> ORGANISM: Danio rerio

<400> SEQUENCE: 28

Met	Leu	Leu	Leu	Leu	Phe	Val	Thr	Val	Cys	Gly	Ser	Ile	Gly	Val	Glu
1					5		10							15	
Ser	Leu	Pro	Leu	Pro	Asp	Ser	Gly	Pro	His	Leu	Ala	Asn	Asp	Trp	Ser
	20					25								30	
Glu	Ala	Val	Arg	Leu	Arg	His	Leu	Tyr	Ala	Ala	Arg	His	Gly	Leu	His
	35					40								45	
Leu	Gln	Ile	Asn	Thr	Asp	Gly	Glu	Ile	Ile	Gly	Ser	Thr	Cys	Lys	Ala
	50					55								60	
Arg	Thr	Val	Ser	Leu	Met	Glu	Ile	Trp	Pro	Val	Asp	Thr	Gly	Cys	Val
	65					70			75					80	
Ala	Ile	Lys	Gly	Val	Ala	Ser	Ser	Arg	Phe	Leu	Cys	Met	Glu	Arg	Leu
	85					90			95						
Gly	Asn	Leu	Tyr	Gly	Ser	His	Ile	Tyr	Thr	Lys	Glu	Asp	Cys	Ser	Phe
	100					105			110						
Leu	Glu	Arg	Ile	Leu	Pro	Asp	Gly	Tyr	Asn	Val	Tyr	Phe	Ser	Ser	Lys
	115					120			125						
His	Gly	Ala	Leu	Val	Thr	Leu	Ser	Gly	Ala	Lys	Asn	Lys	Leu	His	Ser
	130					135			140						
Asn	Asp	Gly	Thr	Ser	Ala	Ser	Gln	Phe	Leu	Pro	Met	Ile	Asn	Thr	Leu
	145					150			155					160	
Ser	Glu	Glu	His	Thr	Lys	Gln	His	Ser	Gly	Glu	Gln	His	Ser	Ser	Val
	165					170			175						
Asn	His	Gly	Gln	Asp	His	Gln	Leu	Gly	Leu	Glu	Ile	Asp	Ser	Met	Asp
	180					185			190						
Pro	Phe	Gly	Lys	Ile	Ser	Gln	Ile	Val	Ile	Gln	Ser	Pro	Ser	Phe	Asn
	195					200			205						
Lys	Arg														
	210														

<210> SEQ_ID NO 29

<211> LENGTH: 215

<212> TYPE: PRT

<213> ORGANISM: Xenopus (Silurana) tropicalis

<400> SEQUENCE: 29

Met	Trp	Lys	Thr	Leu	Pro	Trp	Ile	Leu	Val	Pro	Met	Met	Val	Ala	Val
1						5			10					15	
Leu	Tyr	Phe	Leu	Gly	Gly	Ala	Glu	Ser	Leu	Pro	Leu	Phe	Asp	Ala	Gly
	20					25			30						
Pro	His	Met	Gln	Asn	Gly	Trp	Gly	Glu	Ser	Ile	Arg	Ile	Arg	His	Leu
	35					40			45						
Tyr	Thr	Ala	Arg	Arg	Phe	Gly	His	Asp	Ser	Tyr	Tyr	Leu	Arg	Ile	His
	50					55			60						
Glu	Asp	Gly	Arg	Val	Asp	Gly	Asp	Arg	Gln	Gln	Ser	Met	His	Ser	Leu
	65					70			75					80	
Leu	Glu	Ile	Arg	Ala	Ile	Ala	Val	Gly	Ile	Val	Ala	Ile	Lys	Gly	Tyr
	85					90			95						
Arg	Ser	Ser	Leu	Tyr	Leu	Cys	Met	Gly	Ser	Glu	Gly	Lys	Leu	Tyr	Gly
	100					105			110						
Met	His	Ser	Tyr	Ser	Gln	Asp	Asp	Cys	Ser	Phe	Glu	Glu	Glu	Leu	Leu
	115					120			125						
Pro	Asp	Gly	Tyr	Asn	Met	Tyr	Lys	Ser	Arg	Lys	His	Gly	Val	Ala	Val

170

US 9,475,856 B2

171

172

-continued

130 135 140

Ser Leu Ser Lys Glu Lys Gln Lys Gln Tyr Lys Gly Lys Gly Tyr
 145 150 155 160

Leu Pro Leu Ser His Phe Leu Pro Val Ile Ser Trp Val Pro Met Glu
 165 170 175

Pro Thr Gly Asp Val Glu Asp Asp Ile Tyr Arg Phe Pro Phe Asn Thr
 180 185 190

Asp Thr Lys Ser Val Ile Asp Ser Leu Asp Thr Leu Gly Leu Met Asp
 195 200 205

Phe Ser Ser Tyr His Lys Lys
 210 215

<210> SEQ ID NO 30

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: Otolemur garnettii

<400> SEQUENCE: 30

Met Pro Ser Gly Leu Arg Gly Arg Val Val Ala Gly Ala Leu Ala Leu
 1 5 10 15

Ala Ser Phe Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp
 20 25 30

Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg
 35 40 45

His Leu Tyr Thr Ala Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg
 50 55 60

Val Arg Thr Asp Gly Ala Val Asp Cys Ala Arg Gly Gln Ser Ala His
 65 70 75 80

Ser Leu Leu Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala Ile Lys
 85 90 95

Gly Val His Ser Ala Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met
 100 105 110

Gln Gly Leu Pro Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu
 115 120 125

Ile Arg Pro Asp Gly Tyr Asn Val Tyr Trp Ser Glu Lys His Arg Leu
 130 135 140

Pro Val Ser Leu Ser Ser Ala Arg Gln Arg Gln Leu Tyr Lys Gly Arg
 145 150 155 160

Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Val Thr Pro
 165 170 175

Ala Glu Pro Gly Asp Leu Arg Asp His Leu Glu Ser Asp Met Phe Ser
 180 185 190

Leu Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Thr Arg
 195 200 205

Leu Gly Val Val Lys Ser Pro Ser Phe Gln Lys
 210 215

<210> SEQ ID NO 31

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: Felis catus

<400> SEQUENCE: 31

Met Arg Ser Ala Pro Ser Gln Cys Ala Val Thr Arg Ala Leu Val Leu
 1 5 10 15

Ala Gly Leu Trp Leu Ala Ala Ala Gly Arg Pro Leu Ala Phe Ser Asp

-continued

20 25 30

Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg
35 40 45

His Leu Tyr Thr Ala Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg
50 55 60

Ile Arg Ala Asp Gly Gly Val Asp Cys Ala Arg Ser Gln Ser Ala His
65 70 75 80

Ser Leu Val Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala Ile Lys
85 90 95

Gly Val His Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met
100 105 110

Gln Gly Leu Leu Gln Tyr Ser Ala Gly Asp Cys Ala Phe Gln Glu Glu
115 120 125

Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg Ser Glu Lys His Arg Leu
130 135 140

Pro Val Ser Leu Ser Ser Ala Ile Gln Arg Gln Leu Tyr Lys Gly Arg
145 150 155 160

Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Gly Ser Pro
165 170 175

Ala Glu Pro Arg Asp Leu Gln Asp His Val Glu Ser Glu Arg Phe Ser
180 185 190

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Thr Lys
195 200 205

Met Gly Leu Val Lys Ser Pro Ser Phe Gln Lys
210 215

<210> SEQ_ID NO 32

<211> LENGTH: 223

<212> TYPE: PRT

<213> ORGANISM: Pelodiscus sinensis

<400> SEQUENCE: 32

Met Trp Arg Ser Leu Cys Lys Ser His Thr Ser Leu Ala Leu Leu Gly
1 5 10 15

Leu Cys Phe Ala Val Val Val Arg Ser Leu Pro Phe Ser Asp Ala Gly
20 25 30

Pro His Val Asn Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg His Leu
35 40 45

Tyr Thr Ala Ser Arg His Gly Leu Phe Asn Tyr Phe Leu Arg Ile Ser
50 55 60

Ser Asp Gly Lys Val Asp Gly Thr Ser Ile Gln Ser Pro His Ser Leu
65 70 75 80

Leu Glu Ile Arg Ala Val Ala Val Arg Thr Val Ala Ile Lys Gly Val
85 90 95

His Ser Ser Arg Tyr Leu Cys Met Glu Glu Asp Gly Lys Leu His Gly
100 105 110

Leu Leu Arg Tyr Ser Thr Glu Asp Cys Ser Phe Glu Glu Ile Arg
115 120 125

Pro Asp Gly Tyr Asn Val Tyr Lys Ser Lys Lys Tyr Gly Ile Ser Val
130 135 140

Ser Leu Ser Ser Ala Lys Gln Arg Gln Gln Phe Lys Gly Lys Asp Phe
145 150 155 160

Leu Pro Leu Ser His Phe Leu Pro Met Ile Asn Thr Val Pro Val Glu
165 170 175

US 9,475,856 B2

175

-continued

Ser Met Asp Phe Gly Glu Tyr Gly Asp Tyr Ser His Thr Phe Glu Ser
 180 185 190

Asp Leu Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly
 195 200 205

Ile Thr Ser Lys Ile Ser Pro Val Lys Ser Pro Ser Phe Gln Lys
 210 215 220

<210> SEQ ID NO 33

<211> LENGTH: 217

<212> TYPE: PRT

<213> ORGANISM: Latimeria chalumnae

<400> SEQUENCE: 33

Met Leu Gln Ala Leu Tyr Asn Leu Cys Thr Ala Leu Val Leu Phe Lys
 1 5 10 15

Leu Pro Phe Ala Met Val Gly Tyr Thr Leu Pro Ser Ala Asn Glu Gly
 20 25 30

Pro His Leu Asn Tyr Asp Trp Gly Glu Ser Val Arg Leu Lys His Leu
 35 40 45

Tyr Thr Ser Ser Lys His Gly Leu Ile Ser Tyr Phe Leu Gln Ile Asn
 50 55 60

Asp Asp Gly Lys Val Asp Gly Thr Thr Arg Ser Cys Tyr Ser Leu
 65 70 75 80

Leu Glu Ile Lys Ser Val Gly Pro Gly Val Leu Ala Ile Lys Gly Ile
 85 90 95

Gln Ser Ser Arg Tyr Leu Cys Val Glu Lys Asp Gly Lys Leu His Gly
 100 105 110

Ser Arg Thr Tyr Ser Ala Asp Asp Cys Ser Phe Lys Glu Asp Ile Leu
 115 120 125

Pro Asp Gly Tyr Thr Ile Tyr Val Ser Lys Lys His Gly Ser Val Val
 130 135 140

Asn Leu Ser Asn His Lys Gln Lys Arg Gln Arg Asn Arg Arg Thr Leu
 145 150 155 160

Pro Pro Phe Ser Gln Phe Leu Pro Leu Met Asp Thr Ile Arg Val Glu
 165 170 175

Cys Met Asn Cys Gly Glu His Cys Asp Asp Asn Leu His Asp Glu Leu
 180 185 190

Glu Thr Gly Leu Ser Met Asp Pro Phe Glu Ser Thr Ser Lys Lys Ser
 195 200 205

Phe Gln Ser Pro Ser Phe His Asn Arg
 210 215

<210> SEQ ID NO 34

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: Mustela putorius furo

<400> SEQUENCE: 34

Met Arg Ser Ala Ala Ser Arg Cys Ala Val Ala Arg Ala Leu Val Leu
 1 5 10 15

Ala Gly Leu Trp Leu Ala Ala Gly Arg Pro Leu Ala Phe Ser Asp
 20 25 30

Ala Gly Pro His Val His Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg
 35 40 45

His Leu Tyr Thr Ala Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg
 50 55 60

176

US 9,475,856 B2

177

178

-continued

Ile Arg Ala Asp Gly Gly Val Asp Cys Ala Arg Gly Gln Ser Ala His
65 70 75 80

Ser Leu Val Glu Ile Arg Ala Val Ala Leu Arg Thr Val Ala Ile Lys
85 90 95

Gly Val Tyr Ser Asp Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met
100 105 110

Gln Gly Leu Pro Gln Tyr Ser Ala Gly Asp Cys Ala Phe Glu Glu
115 120 125

Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg Ser Lys Lys His Arg Leu
130 135 140

Pro Val Ser Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asp Arg
145 150 155 160

Gly Phe Leu Pro Leu Ser His Phe Leu Pro Met Leu Pro Gly Ser Leu
165 170 175

Ala Glu Pro Arg Asp Leu Gln Asp His Val Glu Ala Asp Gly Phe Ser
180 185 190

Ala Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Ala Thr Lys
195 200 205

Met Gly Leu Val Lys Ser Pro Ser Phe Gln Lys
210 215

<210> SEQ_ID NO 35

<211> LENGTH: 218

<212> TYPE: PRT

<213> ORGANISM: Takifugu rubripes

<400> SEQUENCE: 35

Ser Ser Thr Arg Ile Ser Gly Asn Met Val Leu Leu Met Leu Pro Ile
1 5 10 15

Thr Val Ala Asn Leu Phe Leu Cys Ala Gly Val Leu Ser Leu Pro Leu
20 25 30

Leu Asp Gln Gly Ser His Phe Pro Gln Gly Trp Glu Gln Val Val Arg
35 40 45

Phe Arg His Leu Tyr Ala Ala Ser Ala Gly Leu His Leu Leu Ile Thr
50 55 60

Glu Glu Gly Ser Ile Gln Gly Ser Ala Asp Pro Thr Leu Tyr Ser Leu
65 70 75 80

Met Glu Ile Arg Pro Val Asp Pro Gly Cys Val Val Ile Arg Gly Ala
85 90 95

Ala Thr Thr Arg Phe Leu Cys Ile Glu Gly Ala Gly Arg Leu Tyr Ser
100 105 110

Ser Gln Thr Tyr Ser Lys Asp Asp Cys Thr Phe Arg Glu Gln Ile Leu
115 120 125

Ala Asp Gly Tyr Ser Val Tyr Arg Ser Val Gly His Gly Ala Leu Val
130 135 140

Ser Leu Gly Asn Tyr Arg Gln Gln Leu Arg Gly Glu Asp Trp Ser Val
145 150 155 160

Pro Thr Leu Ala Gln Phe Leu Pro Arg Ile Ser Ser Leu Asp Gln Asp
165 170 175

Phe Lys Ala Ala Leu Asp Glu Thr Glu Lys Pro Glu Gln Thr Ala Pro
180 185 190

Gln Arg Ser Glu Pro Val Asp Met Val Asp Ser Phe Gly Lys Leu Ser
195 200 205

Gln Ile Ile His Ser Pro Ser Phe His Lys
210 215

-continued

<210> SEQ ID NO 36
<211> LENGTH: 94
<212> TYPE: PRT
<213> ORGANISM: Equus caballus

<400> SEQUENCE: 36

Ala Ala Gly Arg Pro Leu Ala Leu Ser Asp Ala Gly Pro His Val His	
1 5	10 15
Tyr Gly Trp Gly Glu Pro Ile Arg Leu Arg His Leu Tyr Thr Ala Gly	
20 25	30
Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ala Asp Gly Ala	
35 40	45
Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Val Glu Ile Arg	
50 55	60
Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His Ser Val Arg	
65 70	75 80
Tyr Leu Cys Met Gly Ala Asp Gly Arg Met Gln Gly Leu Val	
85	90

<210> SEQ ID NO 37
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Oryzias latipes

<400> SEQUENCE: 37

Thr Met Leu Leu Ile Val Val Thr Ile Ser Thr Met Val Phe Ser Asp	
1 5	10 15
Ser Gly Val Ser Ser Met Pro Leu Ser Asp His Gly Pro His Ile Thr	
20 25	30
His Ser Trp Ser Gln Val Val Arg Leu Arg His Leu Tyr Ala Val Lys	
35 40	45
Pro Gly Gln His Val Gln Ile Arg Glu Asp Gly His Ile His Gly Ser	
50 55	60
Ala Glu Gln Thr Leu Asn Ser Leu Leu Glu Ile Arg Pro Val Ala Pro	
65 70	75 80
Gly Arg Val Val Phe Arg Gly Val Ala Thr Ser Arg Phe Leu Cys Met	
85 90	95
Glu Ser Asp Gly Arg Leu Phe Ser Ser His Thr Phe Asp Lys Asp Asn	
100 105	110
Cys Val Phe Arg Glu Gln Ile Leu Ala Asp Gly Tyr Asn Ile Tyr Ile	
115 120	125
Ser Asp Gln His Gly Thr Leu Leu Ser Leu Gly Asn His Arg Gln Arg	
130 135	140
Gln Gln Gly Leu Asp Arg Asp Val Pro Ala Leu Ala Gln Phe Leu Pro	
145 150	155 160
Arg Ile Ser Thr Leu Gln Gln Gly Val Tyr Pro Val Pro Asp Pro Pro	
165 170	175
His Gln Met Arg Thr Met Gln Thr Glu Lys Thr Leu Asp Ala Thr Asp	
180 185	190
Thr Phe Gly Gln Leu Ser Lys Ile Ile His Ser Pro Ser Phe Asn Lys	
195 200	205

Arg

<210> SEQ ID NO 38
<211> LENGTH: 207

US 9,475,856 B2

181**182**

-continued

<212> TYPE: PRT

<213> ORGANISM: Xiphophorus maculates

<400> SEQUENCE: 38

Met	Phe	Val	Phe	Ile	Leu	Cys	Ile	Ala	Gly	Glu	Leu	Phe	Thr	Leu	Gly
1				5			10					15			
Val	Phe	Cys	Met	Pro	Met	Met	Asp	Gln	Gly	Pro	Leu	Val	Thr	His	Gly
	20				25							30			
Trp	Gly	Gln	Val	Val	Arg	His	Arg	His	Leu	Tyr	Ala	Ala	Lys	Pro	Gly
	35				40							45			
Leu	His	Leu	Leu	Ile	Ser	Glu	Asp	Gly	Gln	Ile	His	Gly	Ser	Ala	Asp
	50				55							60			
Gln	Thr	Leu	Tyr	Ser	Leu	Leu	Glu	Ile	Gln	Pro	Val	Gly	Pro	Gly	Arg
	65				70						75				80
Val	Val	Ile	Lys	Gly	Val	Ala	Thr	Thr	Arg	Phe	Leu	Cys	Met	Glu	Ser
	85				90							95			
Asp	Gly	Arg	Leu	Tyr	Ser	Thr	Glu	Thr	Tyr	Ser	Arg	Ala	Asp	Cys	Thr
	100				105							110			
Phe	Arg	Glu	Gln	Ile	Gln	Ala	Asp	Gly	Tyr	Asn	Val	Tyr	Thr	Ser	Asp
	115				120							125			
Ser	His	Gly	Ala	Leu	Leu	Ser	Leu	Gly	Asn	Asn	Gln	Gln	Arg	His	Ser
	130				135							140			
Gly	Ser	Asp	Arg	Gly	Val	Pro	Ala	Leu	Ala	Arg	Phe	Leu	Pro	Arg	Leu
	145				150						155				160
Asn	Thr	Leu	Gln	Gln	Ala	Val	Pro	Thr	Glu	Pro	Asp	Val	Pro	Asp	Gln
	165				170						175				
Leu	Ser	Pro	Glu	Lys	Val	Gln	Gln	Thr	Val	Asp	Met	Val	Ala	Ser	Phe
	180				185						190				
Gly	Lys	Leu	Ser	His	Ile	Ile	His	Ser	Pro	Ser	Phe	His	Lys	Arg	
	195				200						205				

<210> SEQ_ID NO 39

<211> LENGTH: 220

<212> TYPE: PRT

<213> ORGANISM: Ictidomys tridecemlineatus

<400> SEQUENCE: 39

Met	Arg	Ser	Ala	Pro	Ser	Gly	Arg	Ala	Leu	Ala	Arg	Ala	Leu	Val	Leu	
1				5				10			15					
Ala	Ser	Leu	Trp	Leu	Ala	Val	Ala	Gly	Arg	Pro	Leu	Ala	Arg	Arg	Ser	
	20			25							30					
Leu	Ala	Leu	Ser	Asp	Gln	Gly	Pro	His	Leu	Tyr	Tyr	Trp	Asp	Gln		
	35			40							45					
Pro	Ile	Arg	Leu	Arg	His	Leu	Tyr	Ala	Ala	Gly	Pro	Tyr	Gly	Phe	Ser	
	50			55							60					
Asn	Cys	Phe	Leu	Arg	Ile	Arg	Thr	Asp	Gly	Ala	Val	Asp	Cys	Glu	Glu	
	65			70							75				80	
Lys	Gln	Ser	Glu	Arg	Ser	Leu	Met	Glu	Ile	Arg	Ala	Val	Ala	Leu	Glu	
	85			90							95					
Thr	Val	Ala	Ile	Lys	Asp	Ile	Asn	Ser	Val	Arg	Tyr	Leu	Cys	Met	Gly	
	100			105							110					
Ala	Asp	Gly	Arg	Ile	Gln	Gly	Leu	Pro	Arg	Tyr	Ser	Glu	Glu	Cys		
	115			120							125					
Thr	Phe	Lys	Glu	Glu	Ile	Ser	Tyr	Asp	Gly	Tyr	Tyr	Asn	Val	Tyr	Arg	Ser
	130			135							140					

US 9,475,856 B2

183**184**

-continued

Gln	Lys	Tyr	His	Leu	Pro	Val	Val	Leu	Ser	Ser	Ala	Lys	Gln	Arg	Gln
145				150				155				160			

Leu	Tyr	Gln	Ser	Lys	Gly	Val	Val	Pro	Leu	Ser	Tyr	Phe	Leu	Pro	Met
	165					170			175						

Leu	Pro	Leu	Ala	Ser	Ala	Glu	Thr	Arg	Asp	Arg	Leu	Glu	Ser	Asp	Val
	180					185			190						

Phe	Ser	Leu	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Met	Ala
	195				200				205						

Ser	Glu	Val	Gly	Leu	Lys	Ser	Pro	Ser	Phe	Gln	Lys				
	210				215				220						

<210> SEQ ID NO 40

<211> LENGTH: 203

<212> TYPE: PRT

<213> ORGANISM: Gasterosteus aculeatus

<400> SEQUENCE: 40

Met	Leu	Leu	Leu	Leu	Val	Pro	Ala	Tyr	Val	Ala	Ser	Val	Phe	Leu	Ala
1					5				10				15		

Leu	Gly	Val	Val	Cys	Leu	Pro	Leu	Thr	Asp	Gln	Gly	Leu	His	Met	Ala
	20				25				30						

Asp	Asp	Trp	Gly	Gln	Ser	Val	Arg	Leu	Lys	His	Leu	Tyr	Ala	Ala	Ser
	35				40				45						

Pro	Gly	Leu	His	Leu	Leu	Ile	Gly	Glu	Asp	Gly	Arg	Ile	Gln	Gly	Ser
	50				55			60							

Ala	Gln	Gln	Ser	Pro	Tyr	Ser	Leu	Leu	Glu	Ile	Ser	Ala	Val	Asp	Pro
65					70				75			80			

Gly	Cys	Val	Val	Ile	Arg	Gly	Val	Ala	Thr	Ala	Arg	Phe	Leu	Cys	Ile
	85				90				95						

Glu	Gly	Asp	Gly	Arg	Leu	Tyr	Ser	Ser	Asp	Thr	Tyr	Ser	Arg	Asp	Asp
	100				105				110						

Cys	Thr	Phe	Arg	Glu	Gln	Ile	Leu	Pro	Asp	Gly	Tyr	Ser	Val	Tyr	Val
	115				120				125						

Ser	His	Gly	His	Gly	Ala	Leu	Leu	Ser	Leu	Gly	Asn	His	Arg	Gln	Arg
	130				135				140						

Leu	Gln	Gly	Arg	Asp	His	Gly	Val	Pro	Ala	Leu	Ala	Gln	Phe	Leu	Pro
145					150				155			160			

Arg	Val	Ser	Thr	Met	Asp	Gln	Ala	Ser	Ala	Pro	Asp	Ala	Pro	Gly	Gln
	165				170				175						

Thr	Ala	Thr	Glu	Thr	Glu	Glu	Pro	Val	Asp	Ser	Phe	Gly	Lys	Leu	Ser
	180				185				190						

Gln	Ile	Ile	His	Ser	Pro	Ser	Phe	His	Glu	Arg					
	195				200										

<210> SEQ ID NO 41

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Oreochromis niloticus

<400> SEQUENCE: 41

Met	Leu	Leu	Leu	Leu	Ile	Val	Ser	Ile	Val	Asn	Met	Leu	Phe	Gly	Val
1					5			10		15					

Gly	Met	Val	Cys	Met	Pro	Leu	Ser	Asp	Asn	Gly	Pro	His	Ile	Ala	His
	20				25				30						

Gly	Trp	Ala	Gln	Val	Val	Arg	Leu	Arg	His	Leu	Tyr	Ala	Thr	Arg	Pro
	35				40				45						

-continued

Gly Met His Leu Leu Ile Ser Glu Gly Gly Gln Ile Arg Gly Ser Ala
50 55 60

Val Gln Thr Leu His Ser Leu Met Glu Ile Arg Pro Val Gly Pro Gly
65 70 75 80

Arg Val Val Ile Arg Gly Val Ala Thr Ala Arg Phe Leu Cys Ile Glu
85 90 95

Asp Asp Gly Thr Leu Tyr Ser Ser His Ala Tyr Ser Arg Glu Asp Cys
100 105 110

Ile Phe Arg Glu Gln Ile Leu Pro Asp Gly Tyr Asn Ile Tyr Ile Ser
115 120 125

Asp Arg His Gly Val Leu Leu Ser Leu Gly Asn His Arg Gln Arg Leu
130 135 140

Gln Gly Leu Asp Arg Gly Asp Pro Ala Leu Ala Gln Phe Leu Pro Arg
145 150 155 160

Ile Ser Thr Leu Asn Gln Ile Pro Ser Pro Gly Ala Asn Ile Gly Asp
165 170 175

His Met Lys Val Ala Lys Thr Glu Glu Pro Val Asp Thr Ile Asp Ser
180 185 190

Phe Gly Lys Phe Ser Gln Ile Ile Asp Ser Pro Ser Phe His Lys Arg
195 200 205

<210> SEQ_ID NO 42
<211> LENGTH: 152
<212> TYPE: PRT
<213> ORGANISM: Meleagris gallopavo

<400> SEQUENCE: 42

Val Gly Asn Gln Ser Pro Gln Ser Ile Leu Glu Ile Thr Ala Val Asp
1 5 10 15

Val Gly Ile Val Ala Ile Lys Gly Leu Phe Ser Gly Arg Tyr Leu Ala
20 25 30

Met Asn Lys Arg Gly Arg Leu Tyr Ala Ser Leu Ser Tyr Ser Ile Glu
35 40 45

Asp Cys Ser Phe Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr
50 55 60

Lys Ser Lys Lys Tyr Gly Ile Ser Val Ser Leu Ser Ser Ala Lys Gln
65 70 75 80

Arg Gln Gln Phe Lys Gly Lys Asp Phe Leu Pro Leu Ser His Phe Leu
85 90 95

Pro Met Ile Asn Thr Val Pro Val Glu Val Thr Asp Phe Gly Glu Tyr
100 105 110

Gly Asp Tyr Ser Gln Ala Phe Glu Pro Glu Val Tyr Ser Ser Pro Leu
115 120 125

Glu Thr Asp Ser Met Asp Pro Phe Gly Ile Thr Ser Lys Leu Ser Pro
130 135 140

Val Lys Ser Pro Ser Phe Gln Lys
145 150

<210> SEQ_ID NO 43
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Papio anubis

<400> SEQUENCE: 43

Met Arg Ser Gly Cys Val Val His Ala Trp Ile Leu Ala Ser Leu
1 5 10 15

-continued

Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Pro
20 25 30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35 40 45

Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Thr
50 55 60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65 70 75 80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85 90 95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Lys Met Gln Gly Leu
100 105 110

Leu Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro
115 120 125

Asp Gly Tyr Asn Val Tyr Arg Ser Gln Lys His Arg Leu Pro Val Ser
130 135 140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu
145 150 155 160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Met Ala Pro Glu Glu Pro
165 170 175

Glu Asp Leu Arg Gly Pro Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
180 185 190

Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
195 200 205

Val Arg Ser Pro Ser Phe Glu Lys
210 215

<210> SEQ ID NO 44

<211> LENGTH: 216

<212> TYPE: PRT

<213> ORGANISM: Saimiri boliviensis boliviensis

<400> SEQUENCE: 44

Met Arg Ser Gly Cys Val Val Val His Ala Trp Ile Leu Ala Gly Leu
1 5 10 15

Trp Leu Ala Val Val Gly Arg Pro Leu Ala Phe Ser Asp Ala Gly Pro
20 25 30

His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg His Leu Tyr
35 40 45

Thr Ser Ser Pro His Gly Leu Ser Ser Cys Phe Leu Arg Ile Arg Ser
50 55 60

Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Leu
65 70 75 80

Glu Ile Lys Ala Val Ala Leu Arg Thr Val Ala Ile Lys Gly Val His
85 90 95

Ser Ser Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Leu Gln Gly Leu
100 105 110

Phe Gln Tyr Ser Glu Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro
115 120 125

Asp Gly Tyr Asn Val Tyr Leu Ser Glu Lys His Arg Leu Pro Val Ser
130 135 140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Lys Arg Gly Phe Leu
145 150 155 160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Arg Ala Pro Glu Glu Pro
165 170 175

US 9,475,856 B2

189**190**

-continued

Asp Asp Leu Arg Gly His Leu Glu Ser Asp Val Phe Ser Ser Pro Leu
 180 185 190

Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
 195 200 205

Val Asn Ser Pro Ser Phe Glu Lys
 210 215

<210> SEQ ID NO 45
 <211> LENGTH: 216
 <212> TYPE: PRT
 <213> ORGANISM: Pteropus alecto

<400> SEQUENCE: 45

Met Arg Ser Pro Cys Ala Val Ala Arg Ala Leu Val Leu Ala Gly Leu
 1 5 10 15

Trp Leu Ala Ser Ala Ala Gly Pro Leu Ala Leu Ser Asp Ala Gly Pro
 20 25 30

His Val His Tyr Gly Trp Gly Glu Ala Ile Arg Leu Arg His Leu Tyr
 35 40 45

Thr Ala Gly Pro His Gly Pro Ser Ser Cys Phe Leu Arg Ile Arg Ala
 50 55 60

Asp Gly Ala Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Val
 65 70 75 80

Glu Ile Arg Ala Val Ala Leu Arg Asn Val Ala Ile Lys Gly Val His
 85 90 95

Ser Val Arg Tyr Leu Cys Met Gly Ala Asp Gly Arg Met Leu Gly Leu
 100 105 110

Leu Gln Tyr Ser Ala Asp Asp Cys Ala Phe Glu Glu Ile Arg Pro
 115 120 125

Asp Gly Tyr Asn Val Tyr His Ser Lys Lys His His Leu Pro Val Ser
 130 135 140

Leu Ser Ser Ala Lys Gln Arg Gln Leu Tyr Lys Asp Arg Gly Phe Leu
 145 150 155 160

Pro Leu Ser His Phe Leu Pro Met Leu Pro Arg Ser Pro Thr Glu Pro
 165 170 175

Glu Asn Phe Glu Asp His Leu Glu Ala Asp Thr Phe Ser Ser Pro Leu
 180 185 190

Glu Thr Asp Asp Met Asp Pro Phe Gly Ile Ala Ser Lys Leu Gly Leu
 195 200 205

Glu Glu Ser Pro Ser Phe Gln Lys
 210 215

<210> SEQ ID NO 46
 <211> LENGTH: 245
 <212> TYPE: PRT
 <213> ORGANISM: Myotis davidii

<400> SEQUENCE: 46

Met Ser Gly Gln Asn Ser Gly Arg His Gly Ser Arg Pro Gly Leu Asp
 1 5 10 15

Glu Glu Pro Glu Pro Gly Pro Leu Glu Leu Arg Ala Leu Gly Ser Thr
 20 25 30

Arg Ala Asp Pro Gln Leu Cys Asp Phe Leu Glu Asn His Phe Leu Gly
 35 40 45

Tyr Thr Cys Leu Glu Leu Asp Ile Cys Leu Ala Thr Tyr Leu Gly Val
 50 55 60

-continued

Ser His Trp Gly Glu Ser Ile Arg Leu Arg His Leu Tyr Thr Ser Gly
65 70 75 80

Pro His Gly Pro Ser Ser Cys Phe Leu Arg Ile Arg Val Asp Gly Ala
85 90 95

Val Asp Cys Ala Arg Gly Gln Ser Ala His Ser Leu Val Glu Ile Arg
100 105 110

Ala Val Ala Leu Arg Lys Val Ala Ile Lys Gly Val His Ser Ala Leu
115 120 125

Tyr Leu Cys Met Glu Gly Asp Gly Arg Met Arg Gly Leu Pro Gln Phe
130 135 140

Ser Pro Glu Asp Cys Ala Phe Glu Glu Ile Arg Pro Asp Gly Tyr
145 150 155 160

Asn Val Tyr Arg Ser Gln Lys His Gln Leu Pro Val Ser Leu Ser Ser
165 170 175

Ala Arg Gln Arg Gln Leu Phe Lys Ala Arg Gly Phe Leu Pro Leu Ser
180 185 190

His Phe Leu Pro Met Leu Pro Ser Ser Pro Ala Glu Pro Val His Arg
195 200 205

Glu Arg Pro Leu Glu Pro Asp Ala Phe Ser Ser Pro Leu Glu Thr Asp
210 215 220

Ser Met Asp Pro Phe Gly Ile Ala Asn Asn Leu Arg Leu Val Lys Ser
225 230 235 240

Pro Ser Phe Gln Lys
245

<210> SEQ ID NO 47
<211> LENGTH: 250
<212> TYPE: PRT
<213> ORGANISM: Tupaia chinensis

<400> SEQUENCE: 47

Met Arg Arg Thr Trp Ser Gly Phe Ala Val Ala Thr Arg Ala Gly Ser
1 5 10 15

Pro Leu Ala Leu Ala Asp Ala Gly Pro His Val Asn Tyr Gly Trp Asp
20 25 30

Glu Ser Ile Arg Leu Arg His Leu Tyr Thr Ala Ser Leu His Gly Ser
35 40 45

Thr Ser Cys Phe Leu Arg Ile Arg Asp Asp Gly Ser Val Gly Cys Ala
50 55 60

Arg Gly Gln Ser Met His Ser Leu Leu Glu Ile Lys Ala Val Ala Leu
65 70 75 80

Gln Thr Val Ala Ile Lys Gly Val Tyr Ser Val Arg Tyr Leu Cys Met
85 90 95

Asp Thr Asp Gly Arg Met Gln Gly Leu Pro Gln Tyr Ser Glu Glu Asp
100 105 110

Cys Thr Phe Glu Glu Glu Ile Arg Ser Asp Gly His Asn Val Tyr Arg
115 120 125

Ser Lys Lys His Gly Leu Pro Val Ser Leu Ser Ser Ala Lys Gln Arg
130 135 140

Gln Leu Tyr Lys Gly Arg Gly Phe Leu Ser Leu Ser His Phe Leu Leu
145 150 155 160

Met Met Pro Lys Thr Ser Ala Gly Pro Gly Asn Pro Arg Asp Gln Arg
165 170 175

Asn Pro Arg Asp Gln Arg Asp Pro Asn Thr Phe Ser Leu Pro Leu Glu

US 9,475,856 B2

193**194**

-continued

180 185 190

Thr Asp Ser Met Asp Pro Phe Gly Met Thr Thr Arg His Gly Leu Leu
 195 200 205

Leu Asp Ser Cys Cys Ala Ser Leu Val Leu Leu Asn Ile Ser Thr Asp
 210 215 220

Gly Glu Phe Ser Pro Tyr Gly Asn Ile Leu Arg Pro Ser Phe Arg Phe
 225 230 235 240

Lys Leu Phe Lys Met Lys Lys Val Thr Asn
 245 250

<210> SEQ_ID NO 48

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Heterocephalus glaber

<400> SEQUENCE: 48

Met Arg Phe Ser Lys Ser Thr Cys Gly Phe Phe Asn His Gln Arg Leu
 1 5 10 15

Gln Ala Leu Trp Leu Ser Leu Ser Ser Val Lys Trp Val Leu Asp Ala
 20 25 30

Ala Val Glu Gly Arg Pro Ile Arg Leu Arg His Leu Tyr Ala Ala Gly
 35 40 45

Pro Tyr Gly Arg Ser Arg Cys Phe Leu Arg Ile His Thr Asp Gly Ala
 50 55 60

Val Asp Cys Val Glu Glu Gln Ser Glu His Cys Leu Leu Glu Ile Arg
 65 70 75 80

Ala Val Ala Leu Glu Thr Val Ala Ile Lys Asp Ile Asn Ser Val Arg
 85 90 95

Tyr Leu Cys Met Gly Pro Asp Gly Arg Met Gln Gly Leu Pro Trp Tyr
 100 105 110

Ser Glu Glu Asp Cys Ala Phe Lys Glu Glu Ile Ser Tyr Pro Gly Tyr
 115 120 125

Ser Val Tyr Arg Ser Gln Lys His His Leu Pro Ile Val Leu Ser Ser
 130 135 140

Val Lys Gln Arg Gln Gln Tyr Gln Ser Lys Gly Val Val Pro Leu Ser
 145 150 155 160

Tyr Phe Leu Pro Met Leu Pro Lys Ala Ser Val Glu Pro Gly Asp Glu
 165 170 175

Glu Glu Ser Ala Phe Ser Leu Pro Leu Lys Thr Asp Ser Met Asp Pro
 180 185 190

Phe Gly Met Ala Ser Glu Ile Gly Leu Ala Lys Ser Pro Ser Phe Gln
 195 200 205

Lys

<210> SEQ_ID NO 49

<211> LENGTH: 13

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: C-terminal Portion of FGF19 of the Chimeric Protein

<220> FEATURE:

<221> NAME/KEY: MISC_FEATURE

<222> LOCATION: (7)..(7)

<223> OTHER INFORMATION: Xaa at position 7 is R or N

<400> SEQUENCE: 49

Thr Gly Leu Glu Ala Val Xaa Ser Pro Ser Phe Glu Lys

-continued

 1 5 10

```

<210> SEQ_ID NO 50
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: C-terminal Portion of FGF19 of the Chimeric
  Protein
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (14)..(14)
<223> OTHER INFORMATION: Xaa at position 14 is R or N

<400> SEQUENCE: 50

Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Xaa Ser Pro
1               5               10               15

Ser Phe Glu Lys
20

```

```

<210> SEQ_ID NO 51
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: C-terminal Portion of FGF19 of the Chimeric
  Protein
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: Xaa at position 3 is M or I
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: Xaa at position 4 is V or A
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (13)..(13)
<223> OTHER INFORMATION: Xaa at position 13 is G or R
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: Xaa at position 19 is M or V
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (42)..(42)
<223> OTHER INFORMATION: Xaa at position 42 is R or N

<400> SEQUENCE: 51

```

```

Leu Pro Xaa Xaa Pro Glu Glu Pro Glu Asp Leu Arg Xaa His Leu Glu
1               5               10               15

Ser Asp Xaa Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20               25               30

Gly Leu Val Thr Gly Leu Glu Ala Val Xaa Ser Pro Ser Phe Glu Lys
35               40               45

```

```

<210> SEQ_ID NO 52
<211> LENGTH: 651
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 52

atgcggagcg ggtgtgtggt ggtccacgta tggatcctgg ccggcctctg gctggccgtg     60
gccggggcgcc ccctcgccctt ctccggacgcg gggccccacg tgcactacgg ctggggcgac     120
ccccatccgccc tgccggcacct gtacacacctc ggccccacg ggctctccag ctgcttcctg     180
ccgcattccgtg ccgcacggcgt cgtggactgc gcgcggggcc agagcgcgca cagtttgctg     240

```

-continued

gagatcaagg cagtcgtct gcccggcgtg gccatcaagg gcgtgcacag cgtgcggtag	300
ctctgcattt ggcggcgacgg caagatgcag gggctgtttc agtactcgga ggaagactgt	360
gttttcgagg aggagatccg cccagatggc tacaatgtgt accgatccga gaagcacccgc	420
ctcccggtct ccctgagcag tgccaaacag cggcagctgt acaagaacag aggcttctt	480
caactctctc atttcctgcc catgtgtcccc atgggtcccg aggagcctga ggacccagg	540
ggccacttgg aatctgacat gttctttcg cccctggaga ccgacagcat ggaccattt	600
gggcttgcata ccggacttgg ggcgggtgagg agtcccagct ttgagaagta a	651

<210> SEQ ID NO 53

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Gorilla

<400> SEQUENCE: 53

atgcggaggcg ggtgtgtgg ggtccacgtc tggatcctgg cccgcctctg gctggccgtg	60
gcggggggcc ccctcgccctt ctccggacgcg gggccccacg tgcactacgg ctggggcgac	120
cccateccgcg tgcggcacct gtacacctcc ggccccacg ggctctccag ctgcttctg	180
cgcateccgtg ccgacggcggt cgtggactgc ggcggggggcc agagcgcgcgca cagtttgc	240
gagatcaagg cagtcgtct gcccggcgtg gccatcaagg gcgtgcacag cgtgcggtag	300
ctctgcattt ggcggcgacgg caagatgcag gggctgtttc agtactcgga ggaagactgt	360
gttttcgagg aggagatccg cccagatggc tacaatgtgt accgatctga gaagcacccgc	420
ctcccggtct ccctgagcag tgccaaacag cggcagctgt acaagaacag aggcttctt	480
ccgctctctc atttcctgcc catgtgtcccc atgggtcccg aggagcctga ggacccagg	540
ggccacttgg aatctgacat gttctttcg cccctggaga ccgacagcat ggaccattt	600
gggcttgcata ccggacttgg ggcgggtgagg agtcccagct ttgagaagta a	651

<210> SEQ ID NO 54

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 54

atgcggaaacg ggtgtgtgg ggtccacgtc tggatcctgg cccgcctctg gctggccgtg	60
gcggggggcc ccctcgccctt ctccggacgcg gggccccacg tgcactactg ctggggcgac	120
cccatcccccc tgcggcacct gtacacctcc ggccccatg ggctctccag ctgcttctg	180
cgcateccctg cgaactgcgt catgaactgc ggcggggggcc agagcgcgcgca cagtttgc	240
gagatcaagg cagtcgtct gcccggcgtg gccatcaagg gcgtgcacag cgtgcggtag	300
ctctgcattt ggcggcgacgg caagatgcag gggctgtttc agtactcgga ggaagactgt	360
gttttcgagg aggagatccg cccagatggc tacaatgtgt accgatccga gaagcacccgc	420
ctcccggtct ccctgagcag tgccaaacag cggcagctgt acaagaacag aggcttctt	480
ccactctctc atttcctgcc catgtgtcccc atgggtcccg aggagcctga ggacccagg	540
ggccacttgg aatctgacat gttctttcg cccctggaga ccgacagcat ggaccattt	600
gggcttgcata ccggacttgg ggcgggtgagg agtcccagct ttgagaagta a	651

<210> SEQ ID NO 55

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Macaca mulatta

US 9,475,856 B2

199**200**

-continued

<400> SEQUENCE: 55

atgaggagcg ggtgtgtgg	60
ggtccacgcc tggatcctgg ccagcctctg gctggccgtg	
gcggggcgcc ccctcgccctt ctcggacgcg gggccccacg tgcactacgg ctggggcgac	120
cccatccgcc tgccggcacct gtacacctcc ggccccatg ggctctccag ctgcttcctg	180
cgcacccgca ccgacggcgt cgtggactgc ggcggggcc aaagcgcgca cagttgctg	240
gagatcaagg cagtagctc gcccggcgtg gccatcaagg gcgtgcacag cgtgcggtag	300
ctctgcgtgg ggcggacgg caagatgcag gggctgtttc agtactcaga ggaagactgt	360
gttttcgagg aggagatccg ccctgtatggc tacaatgttat accgatccga gaagcaccgc	420
ctcccggtct ctctgagcag tgccaaacag aggcagctgt acaagaacag aggctttctt	480
cgcctctctc atttcttacc catgtgtccc atggcccccag aggacccatg ggacccagg	540
ggccacttgg aatctgacat gttctttcg cccctggaga ctgacagcat ggaccattt	600
gggcttgtca ccggacttgg ggcgggtgagg agtcccagct ttgagaaata a	651

<210> SEQ ID NO 56

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Pongo abelii

<400> SEQUENCE: 56

atgcggagcg ggtgtgtgg	60
ggtccacgcc tggatcctgg ccggcctctg gctggccgtg	
gcggggcgcc ccctcgccctt ctcggactcg gggccccacg tgcactacgg ctggggcgac	120
cccatccgcc tgccggcacct gtacacctcc ggccccatg ggctctccag ctgcttcctg	180
cgcacccggt cgcacggcgt cgtggactgc ggcggggcc agagcgcgca cagttgctg	240
gagatcaagg cagtagctc gcccggcgtg gccatcaagg gcgtgcacag cgtgcggtag	300
ctctgcgtgg ggcggacgg caagatgcag gggctgtttc agtactcggaa ggaagactgt	360
gttttcgagg aggagatccg cccagatggc tacaatgtgtt accgatccga gaagcaccgc	420
ctcccggtct ccctgagcag tgccaaacag cggcagctgt acaagaacag ggccctttctt	480
cgcctctctc atttcttacc catgtgtccc atggcccccag aggacccatg ggacccagg	540
ggccacttgg aatccgacat gttctttcg cccctggaga ccgacagcat ggaccattt	600
gggcttgtca ccggacttgg ggcgggtgagg agtcccagct ttgagaaata a	651

<210> SEQ ID NO 57

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Nomascus leucogenys

<400> SEQUENCE: 57

atgcggagcg agtgtgtgg	60
ggtccacgcc tggatcctgg ccggcctctg gctggcagtg	
gcggggcgcc ccctcgccctt ttccggacgcg gggccccacg tgcactacgg ctggggcgac	120
cccatccggtc tgccggcacct gtacacctcc ggccccatg ggctctccag ctgcttcctg	180
cgcacccggt cgcacggcgt cgtggactgc ggcggggcc agagcgcgca cagttgctg	240
gagatcaagg cagtagctc gcccggcgtg gccataaagg gcgtgcacag cgtgcggtag	300
ctctgcgtgg ggcggacgg caagatgcag gggctgtttc agtattcggaa ggaagactgt	360
gttttcgagg aggagatccg cccagatggc tacaatgtgtt accgatccga gaagcaccgc	420
ctcccggtct ccctgagcag tgccaaacag cggcagctgt ataagaacag aggctttctt	480

-continued

ccactctctc atttctgtcc catgctgccc atggtcccag aggagcctga ggacctcagg	540
ggccacttgtt aatctgacat gttcttctcg cccctggaga ccgacagcat ggaccattt	600
gggcttgtca ccggactgga ggccgtgagg agtcccagct ttgagaaata a	651
<210> SEQ ID NO 58	
<211> LENGTH: 429	
<212> TYPE: DNA	
<213> ORGANISM: Callithrix jacchus	
<400> SEQUENCE: 58	
atgtgaaagg ccacccgtgg tggccagcag ggacagtccg aagcacaaat gtccacatgt	60
ccccatgttc ctgcgtctct gtggattgtc cagagctgcc tgttttctct gcagctccag	120
tactcggagg aagactgtgc tttcgaggag gagatccgccc ctgatggcta caatgtgtac	180
tggtccgaga agcaccgcct cccggctcc ctgagcagcg ccaaacagcg gcagctgtac	240
aagaaaacgag gctttttcc actgtcccat ttctgtccca tgctgccccat agccccagaa	300
gagcctgagg acctcagggg acacctgaa tctgacgtgt tctcttcacc cctggagact	360
gacagcatgg accccatttg gcttgcacg ggactggagg cggtgaacag tcccagctt	420
gagaagtaa	429
<210> SEQ ID NO 59	
<211> LENGTH: 660	
<212> TYPE: DNA	
<213> ORGANISM: Microcebus murinus	
<400> SEQUENCE: 59	
atgcccggcg ggcaaagcgg ttgtgtggcg gcccgcgcc tgatcctggc cggcctctgg	60
ctgaccggcg cccggggccc gctggccttc tccgacgcgg gcccgcacgt gcactacggc	120
tggggcgagc ccataccgcct cgggcacctg tacaccgcgg gccccacgg cctctccagc	180
tgtttctgc gcataccgcgc agacggctcc gtggactgcgg cgccggggcca gagcgcacac	240
agtttgcgtgg agatcagggc ggtcgcttt cggactgtgg ccatcaaggg cgtgcacagc	300
gtgcggtacc tctgcgtggg cgcagacggc aggtatgcagg ggctgtcccg gtactcggag	360
gaagactgtg cttcgagga ggagatccgc cccgatggct acaacgtgt a cccgtctgag	420
aagcaccgc tgcgggtgtc tctgagcgc gcccggcaga ggcagctgt a caagggcagg	480
ggcttctgc cgctctctca cttoctgtcc atgctgcccc tgacccggc agagaccgg	540
gacctcaggg accacttgg a tccgacatg ttcgcttcgc ccctggagac cgacagcatg	600
gacccttgg ggatcgccac cagacttggg gtggtaaga g tcccagctt tcagaaatga	660
<210> SEQ ID NO 60	
<211> LENGTH: 415	
<212> TYPE: DNA	
<213> ORGANISM: Choloepus hoffmanni	
<400> SEQUENCE: 60	
ttgctcgaaa tgaaggcagt ggcgtgcgg gccgtggcca tcaagggcgt gcacagtgt	60
ctgtacctct gcatgaacgc cgacggcagt ctgcacggc tgcctcggt a ctctgcagaa	120
gactgtgtt ttgaggagga aatccgcccc gacggctaca atgtgtactg gtcttaggaag	180
cacggcctcc ctgtctcttt gacgtgtca aaacagaggc agctgtacaa aggccaggc	240
tttctgcccc tgcctccactt cctgccccatg ctgccccatg cgcggccga gcccgcagac	300
cccgccggatc acgtggagtc ggacatgttc tttcacctc tggaaaccga cagcatggat	360

-continued

ccttttggaa ttgcctccag acttgagctt gtgaacagtc cagcttcag cataa 415

<210> SEQ ID NO 61

<211> LENGTH: 619

<212> TYPE: DNA

<213> ORGANISM: Ailuropoda melanoleuca

<400> SEQUENCE: 61

ggtcctagcc ggcctctgcc tggcggttagc cggggcccccc ctagccttct cggacgcggg	60
gcccacgtg cactacggct ggggtgagcc catccgccta cggcacctgt acacccgcgg	120
cccccacggc ctctccagct gcttcctgccc catccgtccc gacggggggg ttgactgcgc	180
ggggggccag agcgegcaca gtttgggtgaa gatcaggcca gtcgctctgc ggaccgtggc	240
catcaagggt gtgcacagcg tccggtagct ctgcattggcc gggacggca ggatgcaagg	300
gtgcctcag tactctgcag gggactgtgc ttgcaggag gagatccgccc ccgacggcta	360
caatgtgtac cggtccaaga agcaccgtct ccccgatctc ctgagcgggtg ccaaacagag	420
gcagctttac aaagacagag gctttctgcc cctgtccac ttcttgccca tgctgcccgg	480
gagcccagca gagcccgagg acctccagga ccatgcggag tggacgggt tttctgcacc	540
cttagaaaca gacagcatgg acccttttgg gatgccacc aaaatgggac tagtgaagag	600
tcggcgttc cagaataaa	619

<210> SEQ ID NO 62

<211> LENGTH: 660

<212> TYPE: DNA

<213> ORGANISM: Sus scrofa

<400> SEQUENCE: 62

atgcggagcg ctccgagccg gtgcgggtg gtccgcgcgg tggcctggc cggcctctgg	60
ctggccgcag cggggcgccc cctagccctc tcggatgcgt ggccgcacgt gcactacggc	120
tggggcgagt cggtcgcct gcccgcaccc tacactgcga gtcccccacgg cgatccacgc	180
tgcttctgc gcatecactc agaegcccc gtggactgcg cgccgggaca gagcgcgcac	240
agtttgcgtt agatcagggc agtgcgcgtg agtaccgtgg cgatcaaggg cgagcgcagc	300
ggccgttacc tctgcgtgg cgccgcacggc aagatgcga ggcagactca gtactcgat	360
gaggactgtg ctgcgtggg ggagatccgc cctgtatggc acaacgtgtt ctggtccaag	420
aaacaccatc tgccgtctc tctgcgtggc gccaggcaga ggcagctgtt caaaggcagg	480
ggcttctgc cgatgcgttccat cttctgcacc atgcgttcca ctctccacgc cgagccggag	540
gacctccagg acccccttcaa gtccgcaccc ttttcttgc ccctggaaac ggacagcatg	600
gaccctttcc ggatgcgcgc caaactggga gcccgttgcgtt ctataaataaa	660

<210> SEQ ID NO 63

<211> LENGTH: 657

<212> TYPE: DNA

<213> ORGANISM: Bos taurus

<400> SEQUENCE: 63

atgcggagcg ctccgagccg gtgcgggtg gcccgcgcgg tggcctggc tggcctctgg	60
ctggccgcag cggggcgccc cctggccctc tcggatgcgt ggccgcacgt gcactacggc	120
tggggcgagt cggatgcgtt gcccgcaccc tataccgcgg gcccgcaggc cctctacgc	180
tgcttctgc gcatecactc cgacggcgcc gtggactgcg cgccgttgc ggcgcgcac	240

-continued

agtttgatgg agatcagggc ggtcgctcg	agcacccgtag ccatcaaggg cgagcgacg	300
gtgctgtacc tctgcgttgg	cgcgcgacggc aagatgcga	360
gaggactgtg cttecgagga ggagatccgt	cctgacggct acaacgtgt	420
aagcaccatc tccccgttctc	cctgagcgc	480
ggcttcctgc cgctgtctca	cttcctgccc atgctgtcca	540
gacctccagg aaccctgaa gcctgattt	tttctgcccc tgaaaacaga	600
ccttccgggc tcgcccaccaa	actgggatcg gtgaagagtc	657

<210> SEQ ID NO 64

<211> LENGTH: 579

<212> TYPE: DNA

<213> ORGANISM: Canis lupus familiaris

<400> SEQUENCE: 64

ctagccttct ccgcacgcggg	gcccgcacgtg cactccttct	60
cgccacacgt acaccgcggg	ccccacggc ctctccacgt	120
gacggggggg tggactgcgc	gcggggccag agcgcgcaca	180
gtcgctctgc ggaccgtggc	catcaaggcgt	240
gccgacggca ggatgcaagg	gctgcctcag tactccgcgc	300
gagatccgtc ccgatggcta	caatgtgtac tggccaaga	360
ctgagtagtg ccaaaccagag	gcagctctac aaggcagg	420
ttcttaccta tcttgcggg	gagcccaaca gagcccagg	480
tctgacgggt ttctgcatac	cctggaaaca gacagcatgg	540
aaaattggac tagtgaagag	tcccaagtcc caaaaataaa	579

<210> SEQ ID NO 65

<211> LENGTH: 657

<212> TYPE: DNA

<213> ORGANISM: Oryctolagus cuniculus

<400> SEQUENCE: 65

atgcgcggcg cgccgagcgg	agggtgcgcgg gccccgcgcgt	60
ctggccggcg ccgcgcgcgg	cttggcccttg tccgacgcgg	120
tggggcgagc ccgtccgcct	gccccacctg tacgccacca	180
tgcttcctgc gtatacgcgc	cgacggcgcc gtggactgcgc	240
agcttgcgtgg agatccgagc	ggtgcgcctg cgccacgtgg	300
tcccgcgtacc tctgcgttgg	cgccgcacggc aggatgcggg	360
gaggactgtg cttccagga	ggagatcagc tccggctaca	420
caccacgtc ccgtgtctct	gagcgtgcgc aacgtacaa	480
ttccctgcccc tctcccaact	cctggccctg ctggccgg	540
ctcgccgacc accctgaagc	cgaccgttcc tccccggccc	600
cccttcggca tggccaccaa	gctcgccggcgt	657

<210> SEQ ID NO 66

<211> LENGTH: 651

<212> TYPE: DNA

<213> ORGANISM: Pteropus vampyrus

<400> SEQUENCE: 66

-continued

atgcggagcc cgtgcgctgt ggcccgccgc ttggtcctgg cccgcctctg gctggcctca	60
gtgcggggcc ccctcgccct ctccggacgcg gggccgcacg tgcactacgg ctggggcgag	120
gcacatccgcc tcgggcacct gtacaccgcg ggccccacg gcccctccag ctgttctcg	180
cgcacatccgcg cggatggggc ggtggactgc ggcggggccg agagcgcgca cagtttggtg	240
gaaatccggg ctgtcgccct gcggAACGTG gctatcaagg gctgtcacag cgtccgatac	300
ctctgcgtgg gagccgacgg caggatgcta gggctgttc agtactccgc tgacgactgc	360
gccttcgagg aggagatccg cccggacggc tacaacgtgt accactccaa gaaggacccac	420
ctcccggtct ctctgagcag tgccaagcag aggcaactgt acaaggacag gggcttctcg	480
ccctgtccc atttcctgcc catgctgccc aggagccga cagagccga gaacttcgaa	540
gaccacttgg aggccgacac gtttctcg cccctggaga cagacgacat ggacccttt	600
gggattgcca gtaatttggg gctggaggaa agtcccagct tccagaagta a	651

<210> SEQ ID NO 67

<211> LENGTH: 660

<212> TYPE: DNA

<213> ORGANISM: Tursiops truncatus

<400> SEQUENCE: 67

atgcggagcg ctccgagccg gtgcgcgtg gcccgcgcgc ttggtcctggc cggcctctgg	60
ctggctgcag cccggcgccc cctagccctc tccggatgcgg gcccgcacgt gcactacggc	120
tggggggagt ccgtccgcct gccgcacctg tacaccgcgg gtcccccaggg cctctccagc	180
tgcttctgc gcatecactc agacggcgcc gtggactgcg cgccgggtca gagcgcgcac	240
agtttcatgg agatcagggc agtgcgtctg agtaccgtgg ccatcaaggg cgaacgcagc	300
gtcctgtacc tctgcgtgg cgccgcacggc aaaatgcagg ggctgagtcgtactcagct	360
gaggactgtg ctttcgagga gggaaatccgt cccggacggc acaacgtgtc ctggtccaa	420
aaacaccacc tcccggtgtc cctcgacgcg gcccaggcgc ggcagctgtt caaaggcagg	480
ggtttctgc cgctgtctca ctteccccc atgcgtgtca ccatccccac agageccgat	540
gaaatccagg accacttggaa gcccggattt tttgcattgc ccctgaaaac agatagcatg	600
gaccatttg ggctcgccac caaactggga gtgggtgaaga gtcccagctt ctataagtaa	660

<210> SEQ ID NO 68

<211> LENGTH: 660

<212> TYPE: DNA

<213> ORGANISM: Myotis lucifugus

<400> SEQUENCE: 68

atgcggaaacgc cgtggagccg acgcgttgtg gcccggaccc ttggtcctggc cggcctcggtt	60
ctggcctcag cccggggggcc cctcggtttt tccggacgtcg gcccgcacgt gcactacggc	120
tggggggagt ccacccgcct gccgcacctg tacaccctcg gccccacgg cccatccagc	180
tgcttctgc gcateccgcgc tgacggcgca gtggactgcg cgccggggccca gagcgcgcac	240
agtttggatgg agatcagggc cgtgcgtctg cggaaatggg ccatcaaggg cgtgcacagc	300
gcccgttacc tctgcgtgg aggccgcacggc aggtatgtgg ggctgcctca gttctcgccc	360
gaggactgtg ctccgagga ggatccgcg cccggacggc acaacgtgtc ccggtccca	420
aaggcaccgcg tgcccggtctc gctgacgtgg gcccggcaga ggcagctgtt caaggccgg	480
ggcttctgc cgctgtccca ctteccccc atgcgtgtca gcaacccgcg gggacccgtt	540

US 9,475,856 B2

209**210**

-continued

ccccgagaga gccccctcgga gccggacgag ttctcttcgc ccctggaaac agacagcatg	600
gacccttttg ggattgccaa caaacctgagg ctgggtgagaa gtcccagctt tcaggaataa	660

<210> SEQ ID NO 69
<211> LENGTH: 558
<212> TYPE: DNA
<213> ORGANISM: *Ornithorhynchus anatinus*

<400> SEQUENCE: 69

atgcttcct gtgtggttt gcctagtcg ctggagatca aggccgtggc cgtgcgcacg	60
gtggccatca aagggttcca catctctcg tacctctgcg tggaaagagga tggaaaact	120
ccatggcac gtctgctgga gatcaaggcg gtggccgtgc gcacgggtggc catcaaagg	180
gtccacagct ctggcacct ctgcgtggaa gaggatggaa aactccatgg gcagattgg	240
tattctgcag aagactgtgc ttttaagag gaaatacgtc cagatggcta caatgtgtat	300
aaatctaaga aatatgggtt tcctgtttct ttaaggcagcg ccaaacaag gcagcaattc	360
aaaggaaagag actttctgca tctttctcg ttcttgccaa tgatcaacac agtgcctgt	420
gagccagcag agtttggggc ctatgccat tactttgaat cagatataatt ttccctcacct	480
ctggaaaactg acagcatggc cccatggaa attgcccata aactgtcccc tggaaagagc	540
cccaagtttc agaaataa	558

<210> SEQ ID NO 70
<211> LENGTH: 639
<212> TYPE: DNA
<213> ORGANISM: *Monodelphis domestica*

<400> SEQUENCE: 70

atggcccaagc tcctggcccc gctcctcacc ctggctgtctc tctggctggc cccgacggcg	60
cgtgcccac cgctgggtggc cgccgggcct cacgtctact acggctgggg ggagccatt	120
cgtctgcggc atctctacac ggccaatcg cacgggtctcg ccagcttctc ttccctccgg	180
atccaccgcg acggccgcgt ggacggcagc cggagtcaga ggcgcgtcag tttgtggag	240
atcaaggccg tagctttcg gatgtggcg atcaaagggtg tccatagctc tcggtacctg	300
tgtatggag acgcgggaa actccaggga tcggtgaggt tctcggccga ggactgcacc	360
ttcgaggagc agattcgccc cgacggctac aacgtgtacc agtccccaa gtacaacctc	420
cccggtctcg tctgcactga caagcagagg cagcaggccc acggcaagga gcacctgccc	480
ctgtcccaact tcctgccccat gatcaatgtt attcccttgg aggccgagga gcccggggc	540
cccgaggatgt tggccggccgc tctggagacg gacagcatgg acccccttcgg cctcacctcc	600
aagctgttgc cggtaagag ccccaagctt cagaaataa	639

<210> SEQ ID NO 71
<211> LENGTH: 663
<212> TYPE: DNA
<213> ORGANISM: *Anolis carolinensis*

<400> SEQUENCE: 71

atgtgtcgcc gggcggtgcc tctgtgggg gcccctctgg gcttggcgcc cgtggccctcc	60
cgcgcctcc cgctcaccga cgccggggccc cacgtcagct acggctgggg ggagccgtc	120
cggctcaggc acctctacac cgccggggggc cagggccctc tcaagccagtt cctccgcata	180
cacgcccacg ggagagtcga cggcgccggc agccagaacc ggcagagtt gctggagatc	240
cgcgcggctc cggtgcgcgc cgtggccctc aaaggcgtgc acagctcccc ctacctctgc	300

-continued

atggaggagg acggccggct ccgcgggatg ctcagatatt ctgcagaaga ctgttccccc	360
gaagaggaga tgcgtccaga tggctacaat atctacaagt caaagaata cgaggttttg	420
gtctccctaa gtaatgccag acaaagacag caattcaaag ggaaagatTT tcttccccc	480
tctcattct tgccgatgtat caaacatgtg ccagtggagt ctgcagactt tggagagtat	540
ggtgacacca ggcagcatta tgaatcgat atttcagtt cacgtcttga aactgacagc	600
atggaccctt ttggcctcac ttcaagaatgt tcatacgtac aaagtcttag ctttggaaa	660
taa	663

<210> SEQ ID NO 72

<211> LENGTH: 541

<212> TYPE: DNA

<213> ORGANISM: Ochotona princeps

<400> SEQUENCE: 72

gtgcggagca ggggagccat ggcccgccgt ctgggtcttag ccactctctg gctggcccg	60
acgggggggc cgctggccctt gtccgacgctg gggccgcacc tgcactacgg ctggggcgag	120
cccatccggcc tgcggcacct gtacgcccacc agcgccccacg ggctctcgca ctgttttg	180
cgcatacgta ccgaaggcac cgtagactgc gagcgccagcc agagcgccgca cactacagta	240
ctcggaggag gactgcgcct tcgaagagga gatcagctct ggctataacg tgtaccgctc	300
caggaggtac cagctgcccgt tgccctggg cagcgccagg cagaggcagc tgcagcggag	360
ccgtgggttc ctgcccctgt cccacttctt gccgggtctg cccgcggcct cggaggaggt	420
ggcggecccc gctgaccacc cgcaaggcaga cccttctcg cccctggaga ccgacagcat	480
ggaccattt ggaatggcca ccaaggggg gctgggtgaag agccccagct tccagaatgt	540
a	541

<210> SEQ ID NO 73

<211> LENGTH: 666

<212> TYPE: DNA

<213> ORGANISM: Cavia porcellus

<400> SEQUENCE: 73

atgtggagtg cgccgagcgg atgtgtgggt atccgcgcctt tggtcttggc tggcctgtgg	60
ctggcggtgg cggggcgccc cctggcccg cggctctctg cgctatctga ccagggggcg	120
cacttgtact acggctggga ccagccgatc cgccttcggc acctgtacgc cgccggcccc	180
tacgccgcgt cgcgctgttt cctgcgcatt cacacggacg gcgccgttggc ctgcgtcgag	240
gaacagacgc agcactgttt gctggagatc agagcagtcg ctctggagac cgtggccatc	300
aaggacataa acagcgtccg gtacctgtgc atggggcccg acggcaggat gcggggcctg	360
ccctggattt cggaggagga ctgtgccttc aaggaagaga tcaagctaccc gggctacagc	420
gtgttaccgct cccagaagca ccacccccc atcgtgtca gcaatgttcaaa gcaaggcag	480
cagtagccaga gcaaggggggt ggtggccctg tcctacttcc tgcccatgtct gccaaggcc	540
tctgtggagc ccacccccc atcgtgtca gcaatgttcaaa gcaaggcag	600
agcatggacc cctttggat ggccagtgag atcgggttgg tgaagatcc cagcttcag	660
aagtaa	666

<210> SEQ ID NO 74

<211> LENGTH: 593

<212> TYPE: DNA

-continued

<213> ORGANISM: *Tupaia belangeri*

<400> SEQUENCE: 74

atgaggagaa	caccgagcgg	gtttgcagtg	gcccgtgtcc	tcttcctggg	cagccttgg	60
ctggcccgacg	ccggggagccc	cttggccctg	tccgacgccc	ggccgcatgt	gaactacggc	120
tgggatgagt	ccatacgcct	gcgacacttg	tacaccgccta	gcccgcacgg	ctccaccagc	180
tgcttcgtc	gcateccgtga	cgacggctca	gtggactcg	cgccccggca	gagtttgcac	240
agtttgcgg	agatcaaggc	agtgcgtttg	cagaccgtgg	ccatcaaagg	cgtgtacagt	300
gtccgcgtacc	tctgcatttga	cgccgcacggc	aggatgcagg	ggctgggtcc	acgaagcacg	360
gcctcccaagt	ctccctgagc	agtgccttgc	agaggcagct	gttaacgggtt	agggggcttc	420
cttcccttcc	ccacttcctg	ctcatgtgg	ccaagacttc	agcaggggct	ggaaacccca	480
gggaccaccc	agggtctaac	actttcttgt	tgcccttgg	aactgtatgc	atggaccat	540
ttgggatgac	caccagacat	gggctggtga	agagtccca	ctttcaaaaac	taa	593

<210> SEQ ID NO 75

<211> LENGTH: 657

<212> TYPE: DNA

<213> ORGANISM: *Rattus norvegicus*

<400> SEQUENCE: 75

atggcgagaa	agtggagatgg	gcgttattgtg	gcccgcagtc	tggcctggc	cactctgtgg	60
ctggccgtgt	ctggccgtcc	cctgggtccag	caatcccagt	ctgtgtcg	tgaagggtcca	120
ctctttctct	atggctgggg	caagattacc	cgcctgcagt	acctgtactc	tgctgggtccc	180
tacgtctcca	actgcgttcc	gcgttatccgg	agtgcacggct	ctgtggactg	cgaggaggac	240
cagaacacaa	gaaatctgtt	ggagttccgc	gcccgttgc	tgaagacaat	tgccatcaag	300
gacgtcagca	gcgtgcggta	cctctgcatg	agcgcgcacg	gcaagatata	cgggctgatt	360
cgctactcgg	aggaagactg	tacccatgg	gaggaaatgg	actgtttgg	ctacaaccag	420
tacaggttcca	tgaagcacca	cctccacatc	atcttcatca	aggccaagcc	cagagacag	480
ctccaggggcc	agaaacccctc	aaactttatc	cccatatttc	accggcttt	ctttgaatcc	540
acggaccacg	tgaggcttaa	aatgttctct	ctgcccctgg	agagcgacag	catggatcc	600
ttcagaatgg	tggaggatgt	ggaccaccta	gtgaagatgc	ccagcttcca	gaaatga	657

<210> SEQ ID NO 76

<211> LENGTH: 657

<212> TYPE: DNA

<213> ORGANISM: *Mus musculus*

<400> SEQUENCE: 76

atggcgagaa	agtggaaacgg	gcgtgcgggt	gcccgcagcc	tggcctggc	cactctgtgg	60
ctggctgtgt	ctggccgtcc	cctgggtccag	caatcccagt	ctgtgtcg	tgaagatcca	120
ctctttctct	acggctgggg	caagattacc	cgcctgcagt	acctgtactc	cgctgggtccc	180
tatgtctcca	actgcgttcc	ccgatccgg	agcgcacggct	ctgtggactg	cgaggaggac	240
caaaacgaac	gaaatttgc	ggaattccgc	gcccgttgc	tgaagacgt	tgccatcaag	300
gacgtcagca	gcgtgcggta	cctctgcatg	agcgcggacg	gcaagatata	cgggctgatt	360
cgctactcgg	aggaagactg	tacccatgg	gaggaaatgg	actgttttag	ctacaaccag	420
tacagatcca	tgaagcacca	tctccatatac	atcttcatcc	aggccaagcc	cagagaacag	480
ctccaggacc	agaaacccctc	aaactttatc	cccggttttc	accgctccctt	ctttgaaacc	540

-continued

ggggaccagc tgaggctcaa aatgttctcc ctgccccctgg agagtgcacag catggatccg	600
ttcaggatgg tggaggatgt agaccaccta gtgaagagtc ccagcttcca gaaatga	657

<210> SEQ ID NO 77
<211> LENGTH: 684
<212> TYPE: DNA
<213> ORGANISM: Gallus gallus

<400> SEQUENCE: 77

atggggccgg cccgccccgc cgcacccggc gctgccatgg cgctgtgtgg gategcggcc	60
gecgccggcg cccgcccggc cctgcgcgtg cccgacgtcg ggggtccgca cgtcaactac	120
ggctgggggg aacccatccg gctgcccac ctactacacc gcccaggcaa gcacgggctc	180
ttagctgtgt tcctgcgcac cggcgccgac ggccgggtgg acgtgtcggt tagccagagc	240
ccgcagagtc ttgtggagat ccgcgcgtg ggggtgcgcac ccgtggccat caagggcgtg	300
cagagctccc gtcacccatcg catggacgag gggggccggc tgcacccggca gtcagctat	360
tccattgagg actgttccct tgaagaggag attcgtccag acggctacaa cgtgtataaa	420
tcaaagaata acgggatatac ggtgtctttg agcagtgcaca aacaaagaca gcaattcaaa	480
ggaaaagatt ttctcccgct gtctcaacttc ttacccatga tcaacactgt gccagtggag	540
gtgacagact ttgggtaaata tgggtgattac agccaggott ttgagccaga ggtctactca	600
tgcctctcg aaacggacag catggatccc tttggatca cttccaaact gtctccagtg	660
aagagcccca gcttcagaa atga	684

<210> SEQ ID NO 78
<211> LENGTH: 714
<212> TYPE: DNA
<213> ORGANISM: Taeniopygia guttata

<400> SEQUENCE: 78

atggttatca taagcaatct atatctgtg cagaacgtat ttatgtgaa tatggggcga	60
gcaccccttc gcgttcacgc tgctcgctct tggccaccc ctgcctccgc gctgcgcgt	120
ccggccggcc acgcggggcc gcacccataaa tacggctggg gagageccat ccggctgcgg	180
cacccatcaca cggccagcaa gcaegggttc ttcaagctgtc ttctgegtat cggcgctgac	240
ggccgggtgg acgcggccgg cagccagagc ccgcagagcc tgctagagat ccgcgcgtg	300
ggccgtgcgc cccgtggccat caaggccgtg cagagctccc ggtacctgtg catggacgag	360
gccccggcgc tgcacccggca gctcaggaat tccactgaag actgttccct tgaggaggag	420
atccggccatc acggctacaa tgtgtataga tctaaaaaac atgaaatatac ggtgtctttg	480
agcagtgcaca aacaaagaca gcaaggtaag gggaaagatt tcctccctt gtctcaacttc	540
ttggccatga tcaacactgt gcccattggag tcagcagact ttgggtaaata tgggtgattac	600
agccaggccct ttggggcaga ggccttctcc tcacccatgg agacggacag catggacccc	660
tttggcatcg ctcaccaact gtcccttagtg aagagcccta gttccaaaaa ctga	714

<210> SEQ ID NO 79
<211> LENGTH: 633
<212> TYPE: DNA
<213> ORGANISM: Danio rerio

<400> SEQUENCE: 79

atgcgttcctt tactctttgt cactgtttgt ggaagtatcg gctgtggagag cctcccggt	60
--	----

-continued

cccgactctg	gtccacattt	ggcaaatgc	tggagtgaag	ccgtccggct	acgacatctg	120
tacgcagcca	gacatggcctt	acatctgca	ataaacacag	acggagaaat	cattggatcc	180
acatgcaaag	ctcggacagt	aagtttgatg	gagatatggc	cggtggacac	aggctgcgt	240
gcccattaagg	gagttgcaag	ctccccgattt	ctttgcatgg	aaagactggg	aaacctgtac	300
ggatcgacaca	tttacactaa	agaggactgc	tcttttttgg	aacgcacatct	tccagacggc	360
tacaacgtct	acttctcgag	caaacadcgga	getcttgta	ctttaagtgg	tgcgaaaaac	420
aagttgcaca	gtaacgatgg	gacttctgca	tcccagttcc	tccccatgtat	caacacactt	480
tcagaggaac	acactaaaca	gcactcaggg	gaacagca	cttctgttaa	ccatggacag	540
gaccatcagt	tgggccttga	aatagacagt	atggaccctt	tcggaaagat	ctctcaaata	600
gtgatccaga	gtccccagtt	caacaaaaga	tga			633

<210> SEQ_ID NO 80

<211> LENGTH: 648

<212> TYPE: DNA

<213> ORGANISM: Xenopus (Silurana) tropicalis

<400> SEQUENCE: 80

atgtggaaga	ccctgcctt	gattttggtt	cccatgtatgg	tggccgtgt	gtatttcctc	60
ggaggggcgg	aaagtctgcc	gctttttgtat	gccggggccgc	acatgcagaa	cggctgggg	120
gagtcgtatca	gaattcggca	cctgtatacg	gccaggaggt	tcgggcacga	cagctactac	180
ctccggatac	acgaggatgg	cagagtgcgt	ggtgacaggc	aacaaagcat	gcacagttt	240
ttggaaatca	gagcaattgc	agtttggaaatt	gttgccatta	aagggtatcg	cagctctgt	300
tacctgtgca	tgggtccga	gggaaaactc	tatggaatgc	acagttactc	ccaggatgt	360
tgtcttttg	aagaggagct	tctccggat	ggataacaaca	tgtataatc	aaggaaacat	420
ggcggttgc	tctccctaag	caaggagaag	cagaagca	aatacaaagg	aaagggtac	480
ctcccggtgt	cccatccct	acccgtgata	agctgggtgc	ccatggagcc	cacccggat	540
gtagaagatg	atatctacag	gtttccattc	aatacggaca	caaaaagtgt	cattgacagc	600
cttgataccc	tgggactaat	ggattttgc	agtttatcaca	agaaatag		648

<210> SEQ_ID NO 81

<211> LENGTH: 660

<212> TYPE: DNA

<213> ORGANISM: Otolemur garnettii

<400> SEQUENCE: 81

atgcccagcg	ggctgagagg	gcgtgtggta	gccccgcgcc	tggccctggc	cagttctgg	60
ctggccgtgg	ccggggccccc	gctggccttc	tcggatgcgg	gccctcacgt	gcactacggc	120
tggggtgagc	ccatccgcct	gcgcacacgt	tacaccgcgg	gccccacgg	cctctccagc	180
tgtttctgc	gcgtacgcac	cgacgggtcg	gtagactgcg	cgcgggccca	gagcgcacac	240
agtttgcgtt	aaatcagggc	cgtgcgttc	cgaccgtgg	ccatcaaagg	cgtgcacac	300
gcgcggtaacc	tctgcattgg	cgccgacggc	aggatgcagg	ggctgcctca	gtactcgag	360
gaagactgtg	cctttgagga	ggagatccgg	ccagacggct	acaacgtcta	ctggtctgag	420
aagcaccggcc	tgccgggtgc	tctgagcagt	gccccgcaga	ggcagctgt	caagggcagg	480
ggctttctgc	cgctctctca	cttctgcggc	atgtgcgtt	tgacccccc	cgagccccgg	540
gacctcagag	accacctgga	atccgacatg	ttctcttgc	ccctggaaac	tgacagcatg	600
gatccatttg	ggatcgccac	cagactgggc	gtggtgaaga	gtccccagtt	tcagaaatga	660

-continued

<210> SEQ_ID NO 82
<211> LENGTH: 660
<212> TYPE: DNA
<213> ORGANISM: *Felis catus*

<400> SEQUENCE: 82

atgcggagcg	cgccgagcca	gtgcgcggta	acccgcgecc	tggctcttagc	cggctctgg	60
ctggcagcag	ccggggcccc	cctagccttc	tcggacgcgg	ggccteacgt	gcactacggc	120
tggggtgagc	ccatccgact	gcccacactg	tacaccgccc	gccccacgg	cctctccagc	180
tgcttcttc	gcatccgagc	cgacgggggg	gttgactgcg	cgccggccca	gagcgcgcac	240
agtttggtgg	agatcagggc	agtcgtctg	cggaccgtgg	ccatcaaggg	cgtgcacagc	300
gtccggtaacc	tctgcattggg	cggcgcacggc	aggatgcga	ggctgtttca	gtactctgct	360
ggggactgtg	ccttccaaga	ggagatccgc	cccgacggct	acaatgtgt	ccgggtccag	420
aagcacccgtc	tccctgtctc	ttttagttagt	gccatacaga	ggcagctgt	caagggcaga	480
gggtttttgc	ccctgtccca	tttcttgccc	atgtgcggcc	gcagccca	agageccagg	540
gacctccagg	accacgtgga	gtcggagagg	ttttcttcac	ccctggaaac	agacagcatg	600
gacccttttgc	ggattgcccac	aaaaatgggg	tttagtgaaga	gtccccagctt	ccaaaagtaa	660

<210> SEQ_ID NO 83
<211> LENGTH: 672
<212> TYPE: DNA
<213> ORGANISM: *Pelodiscus sinensis*

<400> SEQUENCE: 83

atgtggagga	gcctgtcaa	atctcacacg	tctctggctc	tgctggact	ctgcgttgcg	60
gtggtegtga	gatctctgcc	tttctcgat	gcagggccac	atgtgaacta	tggctgggg	120
gagcctattc	gattaaggca	cctatacacc	gccagcagac	acgggctgtt	caattactc	180
ctgaggatca	gcagtgtatgg	caaagtggat	ggcaccagca	tccagagtcc	tcacagtc	240
ctggaaatca	gggctgtggc	atgtgcac	gtggcgatca	agggcgatca	cagttcccg	300
tacctctgca	ttgaaagaaga	cgggaagctg	catggacttc	tccaggatcc	tacagaagat	360
tgctcccttg	aagaggagat	acgeccagat	ggctacaatg	tatataatc	aaagaatata	420
ggaatctctg	tgtccctaag	tagtgccaaa	caaagacaac	aattcaaaagg	aaaagacttt	480
cttccattgt	ctcacttctt	gcctatgtac	aatacagtac	ctgtggagtc	aatggattt	540
ggagaatatg	gtgattatag	tcatactttt	gaatcagatc	tattctcttc	acctctcgaa	600
actgacagca	tggatccctt	tggaaatcacc	tctaaaatat	ctccagtgaa	gagccccagc	660
tttcagaaat	aa					672

<210> SEQ_ID NO 84
<211> LENGTH: 654
<212> TYPE: DNA
<213> ORGANISM: *Latimeria chalumnae*

<400> SEQUENCE: 84

atgttacagg	cactgtacaa	tctctgtaca	gctctagttt	tgtttaagct	tcctttgca	60
atgggtgggt	acaccctgcc	ttctgccaat	gaagggcccc	atctgaacta	tgactggga	120
gaatctgtaa	gactcaaaca	tctgtacaca	tctagcaac	atggattgtat	cagttactt	180
ttacagatca	atgtatgtgg	caaagttagat	gggaccacta	cacgaagctg	ttatagttt	240

-continued

ctcgaataa aatcagtggg gccaggagtt ttggcaatta aaggcataca gagctccaga	300
tacctttgtg tcgagaagga tggaaaattt catggatcg gcacttattc agcagacat	360
tgctccttca aagaggatat actcccagat ggttacacta tctacgtgc aaagaaacat	420
ggatctgtt gtaatctgag caaccacaaa cagaaacgtc agagaaatcg cagaacctg	480
cctccatTTT ctcaGTTcCT accGCTTATG gacaccATTc gtGTGGAGTG catGAACtGc	540
ggggagcact gtgacgacaa CCTGcatGac gagCTAGAAA caggACTGTC catGGatCC	600
tttgaagta catccaaaaa atcTTTcAG agtcccAGct ttcacaatAG atAA	654

<210> SEQ ID NO 85

<211> LENGTH: 660

<212> TYPE: DNA

<213> ORGANISM: Mustela putorius

<400> SEQUENCE: 85

atgcggagcg ccgcgagtcg gtgcgcggta gcccgcgcgc tggccttagc cgcccttgg	60
ctggcccgac ccggggcccc cctagccctc tccggacgcgg ggccgcacgt gcaactatggc	120
tggggtgagc ccateccgcct acggcacctg tacaccgccc gcccccacgg cctctccagc	180
tgcTTCTGC GcAtCCGTGc CGAcGGCGGG GTTGACTGCG CGCGGGGCCA GAGCGCGCAC	240
agtttggtgg agatccgggc agtcgctctg cggacgggtgg ccatcaaggg cgtgtacagc	300
gaccgcatac tctgcatggg tgccggacggc aggatgcaag ggctgcctca gtactccgCc	360
ggagactgtg CTTCGAGGA GGAGATCCGC CCTGATGGCT ACAACGTGTA CGGGTCCAAG	420
aagcaccgtc tccccgtctc cctgagcagt gcaaacaAAA ggcagctgta caaggaccgg	480
ggcttttgc ctctgtccca ttcttgccc atgctgccc ggagcctggc ggagcccagg	540
gacCTCCAGG accacgtgga ggctgatggg ttTCTGCCC CCCTAGAAAC agacagcatg	600
gaccctttg ggattgccac caaaatggga ctagtgaaga gtcccagctt ccaaaaatga	660

<210> SEQ ID NO 86

<211> LENGTH: 654

<212> TYPE: DNA

<213> ORGANISM: Takifugu rubripes

<400> SEQUENCE: 86

tcatctacaa ggattagtgg aaacatggtt ctccatgc tccccatcac cgTTGCAAAC	60
ctcttcctct gtgctggagt tctctcTTG cTTTGTgg atcaagggtc tcattttccc	120
caaggctggg aacaggtagt ccgcttcagg cacctgtatG ctGCCAGTGC agggctgcac	180
ctgctgatca ctgaagaggg ctcgatccaa ggctctgcag atccaacttt atacagcctg	240
atggagatcc gtccggTgg cccaggctgt gttgtcatta gaggagcagc aaccacacgc	300
ttcctctgca tagaaggTgc tggaagactg tactcatcac agacctacag caaagacgac	360
tgtaccttca gagagcaaAT CCTAGCAGAC ggctacagcg totacagatc tgcggacac	420
ggagctctgg tcagTCTGGG aaactaccgg cagcagctga ggggggagga ctggagcgTT	480
ccgacactgg ctcagttcct ccccaGAAATA agttcaCTgg atcaggactt taaagctgct	540
cttgacgaga ctgagaAGCC agaacaAAACT gcacctcaAA gatcggAAcc tgcgacatg	600
gtggactcat ttggaaagct ctctcagatc atccacagtc ccagTTTCA caag	654

<210> SEQ ID NO 87

<211> LENGTH: 282

<212> TYPE: DNA

<213> ORGANISM: Equus caballus

-continued

<400> SEQUENCE: 87

```

gccccccggc gcccccttagc cttgtccgac gctggggccgc acgtgacta cggctggggc      60
gagcccgatcc gcctgcccga cctgtacacc gccggccccc acggccttc cagctgttc      120
ctgcgcatcc gcgccgatgg cgccgtggac tgcgcggggg gccagagcgc gcacagttt      180
gtggagatca gaggcgtcgc tctgcgcacc gtggccatca agggcgtgca cagcgtccgg      240
tacctctgca tggggccga cggcaggatg caagggtctgg ta                           282

```

<210> SEQ ID NO 88

<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Oryzias latipes

<400> SEQUENCE: 88

```

accatgtgc tcattgttgt caccatttcc acaatggtgt tttctgactc tggagttcc      60
agcatggcgc tctctgtatca tggacccac atcaactcaca gctggagcca agtggtcgc      120
ctccggcacc tgcgtcggt caagcctgga caacatgtcc agatcagaga ggatggacac      180
atccacggct cagcagaaca aactctgaac agcctgtgg agatccgtcc gggtgtcccg      240
ggacgggtgg ttccatcagg agtagccacc tcaagggttc tgcgtatgg gggcggc          300
agactttct cctcacacac atttgacaag gacaactgcg tttccatcaga gcatgttc      360
gcagacggct acaacatcta catttcagat cagcatggaa ccctgtttag tttggaaac      420
caccggcaaa ggcagcaggg tttagaccgg gatgttccag ccctggctca gttccccc      480
aggatcagca ccctgcagca gggcgtgtac ccagtgcac acccccccca ccagatgaga      540
acaatgcaaa cagagaagac tctagatgcc acggacacat ttgggcaact ctctaaaatc      600
attcacagtc ccagttcaa caaaagatga                               630

```

<210> SEQ ID NO 89

<211> LENGTH: 624
<212> TYPE: DNA
<213> ORGANISM: Xiphophorus maculatus

<400> SEQUENCE: 89

```

atgttttgtc tcattctatg cattgttgtt gaactttttt ctctggggagt attttgcatt      60
ccaatgtatgg accagggggcc acttgtcacc catggctggg ggcagggttgt ccggcacccg      120
catctgtatg cagccaaagcc aggactgcac ctactgtatca gtgaggatgg acaaattccac      180
ggttccgcag atcaaactct ttacagcctg ctggagatcc aacctgttgg cccggacgt      240
gttgcgtatca aaggagtggc aaccacacgc ttccctctgca tggagagcga cggcagattg      300
tactcaactg aaacatacag cagagctgac tgacacccatc gagaacagat ccaggcagac      360
ggctacaacg tctcacaccc ttagatccat ggagccctcc tcagtttggg aaacaaccag      420
caaagacaca gcccgtcaga ccgtgggttt ccagctctgg cccgcttct tcccaggatca      480
aacaccccttc agcaggccgt ccccacagag ccggatgttc ctgatcagct cagtcacag     540
aaagtacaac agactgtgga catggtggcc tcctttggca agctctctca tataattcac      600
agtccccagct tccataagag atga                               624

```

<210> SEQ ID NO 90

<211> LENGTH: 663
<212> TYPE: DNA
<213> ORGANISM: Ictidomys tridecemlineatus

-continued

<400> SEQUENCE: 90

atgcggagcg	cggcggagccg	acgtgcctta	gcccgcgc	tgggtgtggc	cagcctctgg	60
ttggcagtgg	ccggacgacc	cctggcccg	cgctctctgg	ctctctccga	ccaggggcca	120
cacttgtact	atggctggga	tcagccccatc	cgcctccggc	acctgtacgc	cgccccccc	180
tacggcttct	ccaactgttt	cctgcgcatac	cgcacccgacg	gcccgtgga	ctgcgaggag	240
aaggcagacg	agcgttagttt	gatggagatc	agggcggtcg	ctctggagac	tgtggccatc	300
aaggacataa	acagegtccg	gtacctctgc	atgggcgc	acggcaggat	acagggactg	360
cctcggta	cggaggaaga	gtgcacgttc	aaggaggaga	tca	cgctacaac	420
gtgttaccgg	cccagaagta	ccacccccc	gtgggtgtca	gcagtgc	ca	480
ctgttaccaga	gcaaggcgt	ggttccctg	tcctacttcc	tgcccatgtc	gcccgtggcc	540
tctgcggaga	ccagggaccg	cttggaaatcc	gatgtgttct	ctttacctct	ggaaactgac	600
agcatggacc	cgtttggat	ggccagtgaa	gtgggcctga	agagccccag	cttccagaag	660
taa						663

<210> SEQ ID NO 91

<211> LENGTH: 612

<212> TYPE: DNA

<213> ORGANISM: Gasterosteus aculeatus

<400> SEQUENCE: 91

atgctgctgc	tgctggtccc	cgcgtacgtt	gccagtgtgt	tttagctct	cgggttgtt	60
tgcttgcccc	taacagatca	gggtctccac	atggccgacg	actggggcca	gtcggtccga	120
ctcaaggcacc	tgtacgcccgc	cagccggga	ctccacctgc	tgatcgaaa	ggatggtcgg	180
atccaaggct	cggcgcagca	aagccctac	agcctgtgg	agatcgtgc	agtggatccg	240
ggctgtgtgg	tcatcagagg	agtagcaacc	gcacggtttc	tctgcacatcga	aggcgatgga	300
agactgtact	catcgacac	ctacagcaga	gacgactgca	cttcaggaa	gcagatctc	360
ccggacggct	acagcgtcta	cgttccat	ggacacgggg	ccctgtctag	cctggggAAC	420
cacaggcaga	ggctgcaggg	tcgagaccac	ggcgtgcggg	ctctggccca	gttcccccgg	480
agggtcagca	ccatggatca	ggcctcgccc	cccga	ccggcagac	cgccaccgg	540
acggaagagc	ccgtggactc	gttggaaag	ctctctcaga	tca	ttcacag	600
cacgagagat	ga					612

<210> SEQ ID NO 92

<211> LENGTH: 627

<212> TYPE: DNA

<213> ORGANISM: Oreochromis niloticus

<400> SEQUENCE: 92

atgctgctgc	tcctcatacg	atccattgtc	aatatgtttt	ttgggtgtgg	aatggtttgc	60	
atgccccctgt	cagacaacgg	gcccacatc	gcccacggct	gggcccagg	ggtccggctc	120	
aggcacccat	acgcccaccag	accggaaatg	cacctgtga	tca	gtgtgggg	180	
cgtgggttctg	ccgtccagac	tctgcacac	cta	atggaga	tccgtccagt	240	
cgtgttgtca	tca	gagggggt	agcaaccgca	agg	ttctctct	300	
ctgtactcat	cgca	ccgtccat	cacagagag	gactgcacat	tca	gagagca	360
gatgggtaca	acatctacat	ctctgacaga	catggagtc	tgctc	atgttgc	420	
cgccaaagac	tgcaggcgtt	agaccgagga	gatccagccc	tggcccagg	cctccccagg	480	

-continued

atcagcactc tgaatcaaat ccctccctt gggcaaaaca tcggtgacca catgaaagta	540
gaaaaaacag aagaacctgt ggacacaata gattcatttg gaaagttctc tcagatcatt	600
gacagtccca gttccataa gagatga	627

<210> SEQ ID NO 93
<211> LENGTH: 456
<212> TYPE: DNA
<213> ORGANISM: Meleagris gallopavo

<400> SEQUENCE: 93

gtaggcaatc aatcaccaca gagcatcattt gaaataactg ctgttgcgtt cggatcgct	60
getatcaagg gcttgcgttc tggcagatac ctggccatga acaaaaagggg caggctttat	120
geatcaactca gctattccat tgaggactgt tcctttaaag aggagatccg tccagatggc	180
tataacgtgt ataaatcaaa gaaatacggg atatcgtgt ctggagcag tgccaaacaa	240
agacaacaat tcaaaaggaaa agatttctc ccactgtctc acttcttacc catgtcaac	300
actgtgcacag tggaggtgac agacttttgtt gaatacgggtt attacagcca ggcttttag	360
ccagaggctt actcatcgcc tctcgaaacg gacagcatgg atccctttgg gatcaactcc	420
aaactgtctc cagtgaagag cccagcttt cagaaa	456

<210> SEQ ID NO 94
<211> LENGTH: 651
<212> TYPE: DNA
<213> ORGANISM: Papio anubis

<400> SEQUENCE: 94

atgaggaggcg ggtgtgtgg ggtccacgcc tggatcctgg ccagcctctg gctggccgt	60
gcggggcgcc ccctcgccctt ctccggacggg gggccccacg tgcactacgg ctggggcgac	120
cccatccggcc tggggcacct gtacacctcc gggccccacg ggctctccag ctgcttcctg	180
cgcacatccgca cccggcggtt cgtggactgc ggcggggggcc aaagcgccca cagttgtg	240
gagatcaagg cagtagctct gggccatgg ggcacatgg ggcgtgcacag cgtgggggtac	300
ctctgcattgg ggcggacgg caagatgcgg gggctgtttc agtactcaga ggaagactgt	360
gttttcggagg aggagatccg ccctgtatggc tacaatgtat accgatccca gaagcaccgc	420
ctccctggctt ccctgagcag tgccaaacag cggcagctgt acaagaacag aggcttttt	480
cgcgtgtctc atttctggcc catgtgtccc atggccccacg aggacgttgg ggacccagg	540
ggcccccttgg aatctgacat gtttcttgg cccctggaga ctgacagcat ggacccattt	600
gggcttgcacccggactggaa ggcgggttggagg agtcccagct ttgagaaata a	651

<210> SEQ ID NO 95
<211> LENGTH: 651
<212> TYPE: DNA
<213> ORGANISM: Saimiri boliviensis boliviensis

<400> SEQUENCE: 95

atgcggaggcg ggtgtgtgg ggtccacgcc tggatcctgg ctggcctctg gctggctgt	60
gtcggggcgcc ccctcgccctt ctccggatgg gggccgcattt tgcattacgg ctggggcgac	120
cccatccggcc tggggcacct gtacacctcc agggccccacg ggctctccag ctgcttcctg	180
cgcacatccgca cccggcggtt cgtggactgc ggcggggggcc agagcgccca cagttgtg	240
gagatcaagg cagtagctct aaggaccgtt ggcacatgg ggcgtgcacag ctcgggggtac	300

-continued

ctctgcatgg	gcccggacgg	caggctgcag	gggctgtcc	agtactcgga	ggaagactgt	360
gcttgcagg	aggagatccg	ccccgacggc	tacaatgtt	acctatccga	gaagcacccgc	420
ctcccggtct	ccctgagcag	cgc当地aaacag	cggcagctgt	acaagaaacg	aggctttctt	480
ccgctgtccc	atttctgcc	catgtgcc	agagccccag	aggagectga	tgacccagg	540
ggccacttgg	aatctgacgt	gttcttca	cccctggaga	ctgatagcat	ggaccattt	600
gggcttgtca	cgggacttgg	ggcggtgaac	agtcccagct	ttgagaagta	a	651

<210> SEQ ID NO 96
<211> LENGTH: 648
<212> TYPE: DNA
<213> ORGANISM: Pteropus alecto

<400> SEQUENCE: 96

atgcgcagcc	cgtgcgcggt	ggcgccgcgc	ctgggtgtgg	cgcccctgtg	gctggcgagc	60
gccccggggcc	cgc当地ggcgct	gagcgatgcg	ggccgcatg	tgcattatgg	ctggggcgaa	120
gccccggggcc	tcgc当地catct	gtataccgcg	ggccgcatg	gccc当地ggcag	ctgctttctg	180
cgc当地tgcg	cgatggcgcc	ggtggattgc	gccc当地ggc	agagcgcc	tagcctgg	240
gaaaattcgcg	cggtggcgct	gccc当地ggc	gccc当地ggc	gccc当地ggc	gccc当地ggc	300
ctgtgcatgg	gcgc当地ggatgg	ccgc当地gtgc	ggcctgctgc	agtatagcgc	ggatgattgc	360
gcgtttgaag	aagaattcgc	cccgatggcc	tataacgtgt	atcatagcaa	aaaacatcat	420
ctgcccgtga	gcctgagcag	cgc当地aacag	cgc当地gtgt	ataaaagatcg	cggcttctg	480
ccgctgagcc	atttctgcc	gatgctgcgc	cgc当地ccgc	ccgatgg	aaacttggaa	540
gatcatctgg	aagcggatac	ctttagcgc	ccgatggaaa	ccgatgat	ggatccgtt	600
ggcattgcga	gcaaaacttgg	cctggaaagaa	agcccgagct	ttcagaaa		648

<210> SEQ ID NO 97
<211> LENGTH: 735
<212> TYPE: DNA
<213> ORGANISM: Myotis davidii

<400> SEQUENCE: 97

atgagcggcc	agaacagcgg	ccgc当地atggc	agccgcccgg	gcctggatga	agaaccggaa	60
ccggggccgc	tggaaactgcg	cgc当地gtggc	agcaccggc	cggatccgc	gctgtgcgat	120
tttctggaaa	accatttct	gggctatacc	tgc当地tggaa	tggatatttgc	cctggcgacc	180
tatctggcg	tgagccatttgc	gggctatacc	tgc当地tggaa	tggatatttgc	cctggcgacc	240
ccgc当地ggcc	c当地gc当地gt	cttctgcgc	attcgc当地gtgg	atggcgccgt	ggattgcgcg	300
cgc当地ggcc	gccc当地ggc	c当地gc当地gt	attcgc当地gtgg	atggcgccgt	ggattgcgcg	360
ataaaaggcg	tgc当地atgcgc	gctgttatctg	tgc当地tggaa	gccc当地ggc	catgegc当地ggc	420
ctgccc当地gt	ttagccgg	agattgcgc	tgc当地tggaa	aaattcgc当地	ggatggctat	480
aacgtgtatc	gcagccagaa	acatcagctg	ccggc当地gtgg	tgagcagcgc	gccc当地ggc	540
cagctgttta	aagcgc当地gg	cttctgcgc	ctgagccattt	ttctgc当地gtgg	gctgc当地ggc	600
agccccggccgg	aaccgggtgca	tgc当地aaacgc	ccgatggaa	c当地gc当地gtgg	tagcgc当地ggc	660
ctggaaaccg	atagcatgg	tccgatggc	attgc当地aa	acctgc当地gtgg	ggtgaaa	720
ccgagcttcc	agaaaa					735

<210> SEQ ID NO 98
<211> LENGTH: 750

231

232

-continued

<212> TYPE: DNA
<213> ORGANISM: *Tupaia chinensis*

<400> SEQUENCE: 98

atgcggccca cctggggcgg ctggcggtg gcgaccgcg cggggcagccc gctggcgctg 60
gccccatgtt gaaactatggc tggatgaaa gcattcgct gcgcacatctg 120
tataccgcga gcctgcatgg cagccaccgc tgctttctgc gcattcgca tgatggcagc 180
gtgggctgca cgccggccca gagcatgcat agccctgctgg aaattaaagc ggtggcgctg 240
cagacccgtgg cgattaaagg cgtgtatagc gtgcgcatac tgcgtcatgga taccgtggc 300
cgccatgcagg gcctgcgcga gtatagcgaa gaagattgca cctttgaaga agaaattcgc 360
agcgtatggcc ataacgtgtt tcgcgcggaa aaacatggcc tgcgggttag cctgagcagc 420
gcgaaaacagc gccagctgtt taaaggccgc ggctttctgaa gcctgagcca tttctgctg 480
atgtatgcga aaaccaggcgc gggcccgggc aacccgcgcg atcagcgca cccggcgat 540
cagcgegatc cgaacacctt tagectggcg ctggaaacccg atagcatgga tccgtttggc 600
atgaccaccc gccatggcct gctgttgat agctgtgcg cgagcttgt gctgtgaac 660
attagcacccg atggcgaatt tagcccgat ggcaacatcc tgccggcagc ctttcgttt 720
aaactgttta aatqaaaaa aqtqaccaac 750

<210> SEQ ID NO 99
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: *Heterocephalus glaber*

<400> SEQUENCE: 99

atgcgcttta gcaaaaacac ctggggcttt tttaccatc agcgctgcga ggcgtgtgg 60
ctgagcctga gcagcgtgaa atgggtgctg gatgcgggg tggaaaggccg cccgattcgc 120
ctgcgcacatc tttatgcggc ggccccgtat ggccgcaccc gctgtttct ggcattcat 180
accgatggcg cggtgattt cgtggaaagaa cagagcgaac attgcctgct gaaaattcgc 240
ggcggtggcgc tggaaaccgt ggcgattaaa gatattaaaca gctgtgcgcta tctgtgcatg 300
ggcccgatg gcccgcattt gggcctggcg tggtatacg aagaagattt cgccgtttaaa 360
aaagaaaattt gctatccggg ctatagcgtg tatcgcagcc agaaacatca tctggcgatt 420
gtgctgagca gctgtaaaca ggcgcagcag tatcagagca aaggcgttgt gccgctgagc 480
tattttctgc cgatgtgcc gaaagcgagc gtggaaaccgg gctgtgtttt gggccggc 540
tttagcctgc cgctgaaac acgcgtatg gatccgtttt gcatggcgag cgaaatttgc 600
ctggcgaaaa gccccggatc ttcaaaaa 627

<210> SEQ ID NO 100
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: *Homo sapiens*

<400> SEQUENCE: 100

```

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
1           5           10          15

```

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
 20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg

-continued

50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65 70 75 80

Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
130 135 140

His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
165 170 175

Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205

Ser

<210> SEQ ID NO 101

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Pongo abelii

<400> SEQUENCE: 101

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65 70 75 80

Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
130 135 140

His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Pro Pro Glu
165 170 175

Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205

Ser

-continued

<210> SEQ_ID NO 102
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 102

Met	Asp	Ser	Asp	Glu	Thr	Gly	Phe	Glu	His	Ser	Gly	Leu	Trp	Val	Ser
1				5				10				15			

Val	Leu	Ala	Gly	Leu	Leu	Leu	Gly	Ala	Cys	Gln	Ala	His	Pro	Ile	Pro
				20				25				30			

Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	Tyr
				35				40			45				

Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His	Leu	Glu	Ile	Arg
				50				55			60				

Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser	Pro	Glu	Ser	Leu
				65			70		75		80				

Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	Ile	Leu	Gly	Val
				85				90			95				

Lys	Thr	Ser	Arg	Phe	Leu	Cys	Gln	Arg	Pro	Asp	Gly	Ala	Leu	Tyr	Gly
				100				105			110				

Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg	Glu	Leu	Leu	Leu
				115			120			125					

Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	Gly	Leu	Pro	Leu
				130		135			140						

His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro	Ala	Pro	Arg	Gly
				145			150		155			160			

Pro	Ala	Arg	Phe	Leu	Pro	Leu	Pro	Gly	Leu	Pro	Pro	Ala	Pro	Pro	Glu
				165			170			175					

Pro	Pro	Gly	Ile	Leu	Ala	Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp
				180			185			190					

Pro	Leu	Ser	Met	Val	Gly	Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Thr
				195			200			205					

Ser

<210> SEQ_ID NO 103
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Canis lupus familiaris

<400> SEQUENCE: 103

Met	Gly	Trp	Ala	Glu	Ala	Gly	Phe	Glu	His	Leu	Gly	Leu	Trp	Val	Pro
1				5				10		15					

Val	Leu	Ala	Val	Leu	Leu	Leu	Glu	Ala	Cys	Arg	Ala	His	Pro	Ile	Pro
				20			25			30					

Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	Tyr
				35			40		45						

Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Glu	Thr	Glu	Ala	His	Leu	Glu	Ile	Arg
				50			55			60					

Ala	Asp	Gly	Thr	Val	Val	Gly	Ala	Ala	Arg	Gln	Ser	Pro	Glu	Ser	Leu
				65			70		75		80				

Leu	Glu	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	Ile	Leu	Gly	Val
				85			90			95					

Lys	Thr	Ser	Arg	Phe	Leu	Cys	Gln	Gly	Pro	Asp	Gly	Thr	Leu	Tyr	Gly
				100			105			110					

-continued

Ser Leu His Phe Asp Pro Val Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Ile Tyr His Ser Glu Thr Leu Gly Leu Pro Leu
130 135 140

Arg Leu Arg Pro His Asn Ser Ala Tyr Arg Asp Leu Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Leu Pro Ala Pro Pro Glu
165 170 175

Pro Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205

Ser

<210> SEQ ID NO 104

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Bos Taurus

<400> SEQUENCE: 104

Met Gly Trp Asp Glu Ala Lys Phe Lys His Leu Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Val Leu Leu Leu Gly Thr Cys Arg Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Ala Asp Gly Thr Val Val Gly Ala Ala Arg Gln Ser Pro Glu Ser Leu
65 70 75 80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Lys Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Pro Lys Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Thr Leu Gly Leu Pro Leu
130 135 140

Arg Leu Pro Pro Gln Arg Ser Ser Asn Arg Asp Pro Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Ala Ala Pro Pro Asp
165 170 175

Pro Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Tyr Gly Arg Ser Pro Ser Tyr Thr
195 200 205

Ser

<210> SEQ ID NO 105

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Equus caballus

<400> SEQUENCE: 105

Met Asp Trp Asp Lys Thr Gly Phe Lys Tyr Gln Gly Leu Trp Val Pro
1 5 10 15

-continued

Val Leu Ala Val Leu Leu Leu Gly Ala Cys Gln Ser His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg His
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Ala Asp Gly Thr Val Ala Gly Ala Val His Arg Ser Pro Glu Ser Leu
65 70 75 80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Thr Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Pro Val Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Thr Leu Gly Leu Pro Leu
130 135 140

Arg Leu Pro His His Ser Ser Pro Tyr Gln Asp Pro Ala Pro Arg Ala
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Phe Pro Pro Ala Pro Pro Glu
165 170 175

Pro Pro Gly Ile Pro Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Arg Ser Arg Ser Pro Ser Tyr Thr
195 200 205

Ser

<210> SEQ ID NO 106

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Ailuropoda melanoleuca

<400> SEQUENCE: 106

Met Gly Trp Asp Glu Ala Arg Ser Glu Gln Leu Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Val Leu Leu Leu Glu Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Ala Ile Arg
50 55 60

Ala Asp Gly Thr Val Val Gly Ala Ala Ser Arg Ser Pro Glu Ser Leu
65 70 75 80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Thr Leu Tyr Gly
100 105 110

Ser Val Arg Phe Asp Pro Val Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Ile Tyr His Ser Glu Thr Leu Gly Leu Pro Leu
130 135 140

Arg Leu Pro Ala His Asn Ser Pro Tyr Arg Asp Ser Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Leu Pro Val Pro Pro Asp
165 170 175

-continued

Pro Pro Gly Ile Leu Gly Pro Glu Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205

Ser

<210> SEQ ID NO 107

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Oryctolagus cuniculus

<400> SEQUENCE: 107

Met Asp Trp Gly Lys Ala Lys Cys Arg Pro Pro Gly Leu Trp Val Pro
1 5 10 15

Ala Leu Ala Ala Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Asp Gln Val Arg Gln Gln His
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Ala Asp Gly Thr Val Val Gly Ala Ala Arg Arg Ser Pro Glu Ser Leu
65 70 75 80

Leu Gln Met Lys Ala Leu Gln Pro Gly Ile Ile Gln Ile Leu Gly Val
85 90 95

Gln Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Thr Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Arg Glu Ala Cys Ser Phe Arg Glu Leu Leu Arg
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Leu Ser Glu Ala Leu Gly Leu Pro Leu
130 135 140

Arg Leu Ser Pro Gly Ser Ser Pro Arg Arg Ala Pro Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Asp Leu Pro Glu
165 170 175

Pro Pro Gly Leu Leu Ala Ala Pro Pro Asp Val Asp Ser Pro Asp
180 185 190

Pro Leu Ser Met Val Gln Pro Ala Leu Asp Gln Ser Pro Ser Tyr Thr
195 200 205

Ser

<210> SEQ ID NO 108

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Gorilla gorilla

<400> SEQUENCE: 108

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu

-continued

65	70	75	80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val			
85	90	95	
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly			
100	105	110	
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu			
115	120	125	
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu			
130	135	140	
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly			
145	150	155	160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Pro Pro Glu			
165	170	175	
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp			
180	185	190	
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala			
195	200	205	

Ser

<210> SEQ_ID NO 109
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Nomascus leucogenys

<400> SEQUENCE: 109

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Pro			
1	5	10	15
Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro			
20	25	30	
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr			
35	40	45	
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg			
50	55	60	
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu			
65	70	75	80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val			
85	90	95	
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly			
100	105	110	
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu			
115	120	125	
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu			
130	135	140	
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly			
145	150	155	160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Pro Pro Glu			
165	170	175	
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp			
180	185	190	
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala			
195	200	205	

Ser

<210> SEQ_ID NO 110

-continued

<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Procavia capensis

<400> SEQUENCE: 110

```

Met Asp Trp Ala Lys Phe Gly Ile Glu His Pro Gly Leu Trp Val Pro
1           5          10          15

Val Met Ala Val Leu Leu Leu Gly Ala Cys Gln Gly Tyr Pro Ile Pro
20          25          30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35          40          45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50          55          60

Ala Asp Gly Thr Val Val Gly Ala Ala His Arg Ser Pro Glu Ser Leu
65          70          75          80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Ile Ile Gln Ile Leu Gly Val
85          90          95

Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Val Leu Tyr Gly
100         105         110

Ser Leu Arg Phe Asp Pro Val Ala Cys Ser Phe Arg Glu Leu Leu Leu
115         120         125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
130         135         140

Arg Leu Pro Ser His Asn Ser Pro Gln Arg Asp Leu Ala Ser Arg Val
145         150         155         160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Arg Leu Thr Val Leu Pro Glu
165         170         175

Pro Ser Gly Val Leu Gly Pro Glu Pro Pro Asp Val Asp Ser Ser Asp
180         185         190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195         200         205

Ser

```

<210> SEQ ID NO 111
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Cavia porcellus

<400> SEQUENCE: 111

```

Met Asp Trp Ala Arg Thr Glu Cys Glu Arg Pro Arg Leu Trp Val Ser
1           5          10          15

Met Leu Ala Ile Leu Leu Val Gly Ala Cys Gln Ala His Pro Ile Pro
20          25          30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35          40          45

Leu Tyr Thr Asp Asp Ala Gln Asp Thr Glu Val His Leu Glu Ile Arg
50          55          60

Ala Asp Gly Ser Val Arg Gly Ile Ala His Arg Ser Pro Glu Ser Leu
65          70          75          80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Ile
85          90          95

Arg Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ser Leu Tyr Gly
100         105         110

Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
115         120         125

```

-continued

Ala Asp Gly Tyr Asn Val Tyr Lys Ser Glu Ala His Gly Leu Pro Leu
 130 135 140

His Leu Leu Arg Gly Asp Ser Leu Ser Gln Glu Pro Ala Pro Pro Gly
 145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Ala Thr Pro Pro Glu
 165 170 175

Pro Pro Arg Met Leu Pro Pro Gly Pro Pro Asp Val Gly Ser Ser Asp
 180 185 190

Pro Leu Ser Met Val Gly Pro Leu Trp Asp Arg Ser Pro Ser Tyr Thr
 195 200 205

Ser

<210> SEQ ID NO 112

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Tupaia belangeri

<400> SEQUENCE: 112

Met Gly Trp Asp Lys Ala Arg Phe Glu His Leu Gly Ala Trp Ala Pro
 1 5 10 15

Val Leu Ala Val Leu Leu Leu Gly Ala Cys Gln Ala Tyr Pro Ile Pro
 20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
 35 40 45

Leu Tyr Thr Asp Asp Thr Gln Asp Thr Glu Ala His Leu Glu Ile Arg
 50 55 60

Ala Asp Gly Thr Val Val Gly Ala Ala His Gln Ser Pro Glu Ser Leu
 65 70 75 80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
 100 105 110

Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
 115 120 125

Glu Asp Gly Tyr Asn Ile Tyr Gln Ser Glu Ala Arg Gly Leu Pro Leu
 130 135 140

Arg Leu Pro Pro His Asp Ser Pro His Arg Asp Arg Arg Thr Pro Arg Gly
 145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Leu Val Pro Pro Glu
 165 170 175

Leu Pro Gly Val Leu Ala Leu Glu Pro Pro Asp Val Gly Ser Ser Asp
 180 185 190

Pro Leu Ser Met Met Gly Pro Ser Gln Gly Gln Ser Pro Ser Tyr Ala
 195 200 205

Ser

<210> SEQ ID NO 113

<211> LENGTH: 210

<212> TYPE: PRT

<213> ORGANISM: Sorex araneus

<400> SEQUENCE: 113

Met Val Trp Asp Lys Ala Arg Gly Gln Gln Leu Gly Leu Trp Ala Pro
 1 5 10 15

Met Leu Leu Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Leu Pro
 20 25 30

-continued

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Leu Arg Phe
 35 40 45
 Leu Tyr Thr Asp Asp Ala Gln Arg Thr Gly Ala His Leu Glu Ile Arg
 50 55 60
 Ala Asp Gly Thr Val Gln Gly Ala Ala His Arg Thr Pro Glu Cys Leu
 65 70 75 80
 Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95
 Ser Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Val Leu Tyr Gly
 100 105 110
 Ser Leu Arg Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
 115 120 125
 Gln Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala Leu Gly Leu Pro Leu
 130 135 140
 Tyr Leu His Pro Pro Ser Ala Pro Val Ser Gln Glu Pro Ala Ser Arg
 145 150 155 160
 Gly Ala Val Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Ser Leu
 165 170 175
 Glu Pro Pro Arg Pro Pro Ala Pro Val Pro Pro Asp Val Gly Ser Ser
 180 185 190
 Asp Pro Leu Ser Met Val Gly Pro Pro Glu Arg His Ser Pro Ser Tyr
 195 200 205
 Thr Ser
 210

<210> SEQ ID NO 114
 <211> LENGTH: 208
 <212> TYPE: PRT
 <213> ORGANISM: Ictidomys tridecemlineatus
 <400> SEQUENCE: 114

Met Asp Trp Val Lys Ala Lys Leu Glu Pro Leu Gly Leu Trp Val Leu
 1 5 10 15
 Val Leu Ala Ala Leu Val Leu Gly Ala Cys Gln Ala Tyr Pro Ile Pro
 20 25 30
 Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
 35 40 45
 Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
 50 55 60
 Ala Asp Gly Thr Val Val Gly Ala Ala His Gln Ser Pro Glu Ser Leu
 65 70 75 80
 Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95
 Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Val Leu Tyr Gly
 100 105 110
 Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Gln Leu Leu
 115 120 125
 Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ser His Gly Leu Pro Val
 130 135 140
 Arg Leu Pro Pro Asn Ser Pro Tyr Arg Asp Pro Ala Pro Pro Gly Pro
 145 150 155 160
 Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Ala Leu Glu Pro
 165 170 175
 Pro Gly Ile Leu Gly Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro

-continued

180

185

190

Leu Ser Met Val Gly Pro Leu Gln Gly Arg Ser Pro Ser Tyr Ala Ser
195 200 205

<210> SEQ ID NO 115

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Loxodonta Africana

<400> SEQUENCE: 115

Met Asp Trp Ala Lys Phe Gly Leu Glu His Pro Gly Leu Trp Val Pro
1 5 10 15

Val Met Ala Val Leu Leu Leu Gly Ala Cys Gln Gly His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Gln Glu Thr Glu Ala His Leu Glu Ile Arg Ala
50 55 60

Asp Gly Thr Val Ala Gly Ala Ala His Arg Ser Ser Glu Ser Leu Leu
65 70 75 80

Glu Leu Lys Ala Leu Lys Pro Gly Ile Ile Gln Ile Leu Gly Val Lys
85 90 95

Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Val Leu Tyr Gly Ser
100 105 110

Leu His Phe Asp Pro Ala Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
115 120 125

Asp Gly Tyr Asn Val Tyr Trp Ser Glu Ala His Gly Leu Pro Ile Arg
130 135 140

Leu Pro Ser His Asn Ser Pro Tyr Arg Asp Pro Ala Ser Arg Val Pro
145 150 155 160

Ala Arg Phe Leu Pro Leu Pro Gly Leu Leu Pro Met Leu Gln Glu Pro
165 170 175

Pro Gly Val Leu Ala Pro Glu Pro Pro Asp Val Asp Ser Ser Asp Pro
180 185 190

Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
195 200 205

<210> SEQ ID NO 116

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Sus scrofa

<400> SEQUENCE: 116

Met Gly Trp Ala Glu Ala Lys Phe Glu Arg Leu Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Val Leu Leu Gly Ala Cys Gln Ala Arg Pro Ile Pro Asp
20 25 30

Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
35 40 45

Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg Ala
50 55 60

Asp Gly Thr Val Ala Gly Val Ala Arg Gln Ser Pro Glu Ser Leu Leu
65 70 75 80

Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Gln
85 90 95

Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Arg Leu Tyr Gly Ser

US 9,475,856 B2

253**254**

-continued

100 105 110

Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
 115 120 125

Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala Leu Gly Leu Pro Leu Arg
 130 135 140

Leu Pro Pro His Arg Ser Ser Asn Arg Asp Leu Ala Pro Arg Gly Pro
 145 150 155 160

Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Pro Pro Glu Pro
 165 170 175

Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro
 180 185 190

Leu Ser Met Val Gly Pro Ser His Gly Arg Ser Pro Ser Tyr Thr Ser
 195 200 205

<210> SEQ ID NO 117

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Felis catus

<400> SEQUENCE: 117

Met Asp Trp Asp Glu Ala Gly Ser Gln Arg Leu Gly Leu Trp Val Val
 1 5 10 15

Leu Gly Val Leu Leu Pro Glu Ala Cys Gln Ala His Pro Ile Pro Asp
 20 25 30

Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Phe Leu
 35 40 45

Tyr Thr Asp Asp Ala Gln Glu Thr Glu Val His Leu Glu Ile Lys Ala
 50 55 60

Asp Gly Thr Val Val Gly Thr Ala Arg Arg Ser Pro Glu Ser Leu Leu
 65 70 75 80

Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys
 85 90 95

Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Thr Leu Tyr Gly Ser
 100 105 110

Leu Arg Phe Asp Pro Ala Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
 115 120 125

Asp Gly Tyr Asn Ile Tyr His Ser Glu Thr Leu Gly Leu Pro Leu Arg
 130 135 140

Leu Pro Pro His Asn Ser Pro Tyr Arg Asp Leu Ala Pro Arg Ala Pro
 145 150 155 160

Ala Arg Phe Leu Pro Leu Pro Gly Leu Leu Pro Ala Pro Pro Glu Pro
 165 170 175

Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro
 180 185 190

Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 195 200 205

<210> SEQ ID NO 118

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Otolemur garnettii

<400> SEQUENCE: 118

Asp Lys Ala Arg Thr Gly Phe Lys His Pro Gly Pro Trp Phe Pro Leu
 1 5 10 15

Leu Ala Val Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro Asp

US 9,475,856 B2

255**256**

-continued

20 25 30

Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
 35 40 45

Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg Glu
 50 55 60

Asp Gly Thr Val Val Gly Ala Ala Gln Gln Ser Pro Glu Ser Leu Leu
 65 70 75 80

Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys
 85 90 95

Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Gly Leu Tyr Gly Ser
 100 105 110

Leu Tyr Phe Asp Pro Lys Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
 115 120 125

Asp Gly Tyr Asn Val Tyr Trp Ser Glu Thr Tyr Gly Leu Pro Leu His
 130 135 140

Leu Pro Pro Ala Asn Ser Pro Tyr Trp Gly Pro Ser Leu Arg Ser Pro
 145 150 155 160

Ala Arg Phe Leu Pro Leu Pro Gly Pro Pro Ala Ala Ser Pro Glu Leu
 165 170 175

Pro Gly Ile Leu Ala Leu Glu Pro Pro Asp Val Gly Ser Ser Asp Pro
 180 185 190

Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 195 200 205

<210> SEQ ID NO 119

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 119

Met Asp Trp Met Lys Ser Arg Val Gly Ala Pro Gly Leu Trp Val Cys
 1 5 10 15

Leu Leu Leu Pro Val Phe Leu Leu Gly Val Cys Glu Ala Tyr Pro Ile
 20 25 30

Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg
 35 40 45

Tyr Leu Tyr Thr Asp Asp Gln Asp Thr Glu Ala His Leu Glu Ile
 50 55 60

Arg Glu Asp Gly Thr Val Val Gly Thr Ala His Arg Ser Pro Glu Ser
 65 70 75 80

Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly
 85 90 95

Val Lys Ala Ser Arg Phe Leu Cys Gln Gln Pro Asp Gly Thr Leu Tyr
 100 105 110

Gly Ser Pro His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu
 115 120 125

Leu Lys Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro
 130 135 140

Leu Arg Leu Pro Gln Lys Asp Ser Gln Asp Pro Ala Thr Arg Gly Pro
 145 150 155 160

Val Arg Phe Leu Pro Met Pro Gly Leu Pro His Glu Pro Gln Glu Gln
 165 170 175

Pro Gly Val Leu Pro Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro
 180 185 190

US 9,475,856 B2

257

-continued

Leu Ser Met Val Glu Pro Leu Gln Gly Arg Ser Pro Ser Tyr Ala Ser	
195	200
	205

<210> SEQ ID NO 120

<211> LENGTH: 210

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 120

Met Glu Trp Met Arg Ser Arg Val Gly Thr Leu Gly Leu Trp Val Arg	
1	5
	10
	15

Leu Leu Leu Ala Val Phe Leu Leu Gly Val Tyr Gln Ala Tyr Pro Ile	
20	25
	30

Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg	
35	40
	45

Tyr Leu Tyr Thr Asp Asp Asp Gln Asp Thr Glu Ala His Leu Glu Ile	
50	55
	60

Arg Glu Asp Gly Thr Val Val Gly Ala Ala His Arg Ser Pro Glu Ser	
65	70
	75
	80

Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly	
85	90
	95

Val Lys Ala Ser Arg Phe Leu Cys Gln Gln Pro Asp Gly Ala Leu Tyr	
100	105
	110

Gly Ser Pro His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu	
115	120
	125

Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro	
130	135
	140

Leu Arg Leu Pro Gln Lys Asp Ser Pro Asn Gln Asp Ala Thr Ser Trp	
145	150
	155
	160

Gly Pro Val Arg Phe Leu Pro Met Pro Gly Leu Leu His Glu Pro Gln	
165	170
	175

Asp Gln Ala Gly Phe Leu Pro Pro Glu Pro Pro Asp Val Gly Ser Ser	
180	185
	190

Asp Pro Leu Ser Met Val Glu Pro Leu Gln Gly Arg Ser Pro Ser Tyr	
195	200
	205

Ala Ser
210

<210> SEQ ID NO 121

<211> LENGTH: 107

<212> TYPE: PRT

<213> ORGANISM: Vicugna pacos

<400> SEQUENCE: 121

Met Asp Trp Asp Glu Ala Lys Phe Glu His Arg Gly Leu Trp Val Pro	
1	5
	10
	15

Val Leu Thr Val Leu Leu Gly Ala Cys Gln Ala Arg Pro Ile Pro	
20	25
	30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr	
35	40
	45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg	
50	55
	60

Ala Asp Gly Thr Val Val Gly Val Ala Arg Gln Pro Glu Gly Ile Pro	
65	70
	75
	80

Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly	
85	90
	95

258

US 9,475,856 B2

259**260**

-continued

Pro Ser Tyr Ser Arg Ser Pro Ser Tyr Thr Ser
100 105

<210> SEQ ID NO 122
<211> LENGTH: 211
<212> TYPE: PRT
<213> ORGANISM: Anolis carolinensis
<400> SEQUENCE: 122

Cys Lys Ser Lys Gly Gly Gly Lys Gly Glu Arg Met Trp Val Asp
1 5 10 15

Leu Val Phe Trp Ala Ala Leu Leu Arg Thr Ala Pro Ala Leu Pro Leu
20 25 30

Arg Asn Ser Asn Pro Ile Tyr Gln Phe Asp Gly Gln Val Arg Leu Arg
35 40 45

His Leu Tyr Thr Ala Asp Glu Gln Thr His Leu His Leu Glu Ile Leu
50 55 60

Pro Asp Gly Thr Val Gly Gly Ser Arg Phe Gln Asn Pro Phe Ser Leu
65 70 75 80

Met Glu Ile Lys Ala Val Lys Pro Gly Val Ile Arg Met Gln Ala Lys
85 90 95

Lys Thr Ser Arg Phe Leu Cys Met Lys Pro Asn Gly Arg Leu Tyr Gly
100 105 110

Ser Leu Phe Tyr Ser Glu Glu Ala Cys Asn Phe His Glu Lys Val Leu
115 120 125

Ser Asp Gly Tyr Asn Leu Tyr Tyr Ser Glu Asn Tyr Asn Ile Pro Val
130 135 140

Ser Leu Ser Ser Ala Gly Asn Leu Gly Gln Ser Arg Gln Leu Pro Pro
145 150 155 160

Phe Ser Gln Phe Leu Pro Leu Val Asn Lys Ile Pro Leu Glu Pro Val
165 170 175

Leu Glu Asp Phe Asp Phe Tyr Gly His Gln Leu Asp Val Glu Ser Ala
180 185 190

Asp Pro Leu Ser Ile Leu Gly Gln Asn Pro Gly Phe Met Ser Pro Ser
195 200 205

Tyr Val Phe
210

<210> SEQ ID NO 123
<211> LENGTH: 188
<212> TYPE: PRT
<213> ORGANISM: Gadus morhua
<400> SEQUENCE: 123

Leu Leu Leu Ala Thr Leu Leu His Ile Gly Leu Ser Phe Tyr Val Pro
1 5 10 15

Asp Ser Gly Pro Leu Leu Trp Leu Gly Asp Gln Val Arg Glu Arg His
20 25 30

Leu Tyr Thr Ala Glu Ser His Arg Arg Gly Leu Phe Leu Glu Met Ser
35 40 45

Pro Asp Gly Gln Val Thr Gly Ser Ala Ala Gln Thr Pro Leu Ser Val
50 55 60

Leu Glu Leu Arg Ser Val Arg Ala Gly Asp Thr Val Ile Arg Ala Arg
65 70 75 80

Leu Ser Ser Leu Tyr Leu Cys Val Asp Arg Ala Gly His Leu Thr Gly
85 90 95

-continued

Gln Arg Gln Tyr Thr Glu Ser Asp Cys Thr Phe Arg Glu Val Ile Leu
100 105 110

Glu Asp Gly Tyr Thr His Phe Leu Ser Val His His Gly Leu Pro Ile
115 120 125

Ser Leu Ala Pro Arg His Ser Pro Gly Arg Gln Gly Leu Arg Phe Ser
130 135 140

Arg Phe Leu Pro Leu Arg Ser Ser Leu Ser Glu Asp Arg Val Ala Glu
145 150 155 160

Pro Pro Asp Ser Pro Leu Asn Leu Asp Ser Glu Asp Pro Leu Gly Met
165 170 175

Gly Leu Gly Ser Leu Leu Ser Pro Ala Phe Ser Met
180 185

<210> SEQ_ID NO 124

<211> LENGTH: 216

<212> TYPE: PRT

<213> ORGANISM: Latimeria chalumnae

<400> SEQUENCE: 124

Met Leu Cys Gln Ser Phe Val Ile Leu Ser Gln Lys Phe Ile Phe Gly
1 5 10 15

Leu Phe Leu Thr Gly Leu Gly Leu Thr Gly Leu Ala Trp Thr Arg Pro
20 25 30

Phe Gln Asp Ser Asn Pro Ile Leu Gln Tyr Ser Asp Ser Ile Arg Leu
35 40 45

Arg His Leu Tyr Thr Ala Ser Glu Ser Arg His Leu His Leu Gln Ile
50 55 60

Asn Ser Asp Gly Gln Val Gly Gly Thr Thr Lys Gln Ser Pro Tyr Ser
65 70 75 80

Leu Leu Glu Met Lys Ala Val Lys Thr Gly Phe Val Val Ile Arg Gly
85 90 95

Lys Lys Ser Ala Arg Tyr Leu Cys Met Glu Arg Ser Gly Arg Leu Tyr
100 105 110

Gly Ser Leu Gln Tyr Thr Glu Lys Asp Cys Thr Phe Lys Glu Val Val
115 120 125

Leu Ala Asp Gly Tyr Asn Leu Tyr Val Ser Glu Glu His Gln Ala Thr
130 135 140

Val Thr Leu Ser Pro Met Arg Ala Arg Ile Ala Gln Gly Lys Lys Ile
145 150 155 160

Pro Pro Phe Ser His Phe Leu Pro Met Val Asn Lys Val Pro Val Glu
165 170 175

Asp Val Ala Ala Glu Met Glu Phe Val Gln Val Leu Arg Glu Met Thr
180 185 190

Ala Asp Val Asp Ser Pro Asp Pro Phe Gly Met Thr Trp Glu Glu Ser
195 200 205

Val His Ser Pro Ser Phe Phe Ala
210 215

<210> SEQ_ID NO 125

<211> LENGTH: 192

<212> TYPE: PRT

<213> ORGANISM: Tursiops truncates

<400> SEQUENCE: 125

Met Gly Trp Asp Lys Thr Lys Leu Glu His Leu Gly Leu Trp Val Pro
1 5 10 15

-continued

Val Leu Ala Val Leu Leu Gly Pro Cys Gln Ala His Pro Ile Pro Asp
20 25 30

Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
35 40 45

Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg Ala
50 55 60

Asp Gly Thr Val Val Gly Thr Ala Arg Arg Ser Pro Glu Gly Val Lys
65 70 75 80

Thr Ser Arg Phe Leu Cys Gln Gly Pro Glu Gly Arg Leu Tyr Gly Ser
85 90 95

Leu His Phe Asn Pro Gln Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
100 105 110

Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala Leu Gly Ile Pro Leu Arg
115 120 125

Leu Pro Pro His Arg Ser Ser Asn Trp Asp Leu Ala Pro Arg Gly Pro
130 135 140

Ala Arg Phe Leu Pro Leu Pro Gly Phe Leu Pro Pro Pro Leu Glu Pro
145 150 155 160

Pro Gly Ile Leu Ala Pro Glu Pro Pro Asn Val Gly Ser Ser Asp Pro
165 170 175

Leu Ser Met Val Gly Pro Ser His Gly Arg Ser Pro Ser Tyr Thr Ser
180 185 190

<210> SEQ ID NO 126

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Mustela putorius furo

<400> SEQUENCE: 126

Met Gly Trp Glu Glu Ala Arg Ser Glu His Leu Gly Leu Trp Val Pro
1 5 10 15Val Leu Ala Val Leu Leu Gly Ala Cys Gln Ala Tyr Pro Ile Pro
20 25 30Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50 55 60Ala Asp Gly Thr Val Val Gly Ala Ala Arg Arg Ser Pro Glu Ser Leu
65 70 75 80Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asn Gly Thr Leu Tyr Gly
100 105 110Ser Phe His Phe Asp Pro Val Ala Cys Ser Phe Arg Glu Val Leu Leu
115 120 125Glu Asp Gly Tyr Asn Ile Tyr His Ser Glu Thr Leu Gly Leu Pro Leu
130 135 140Arg Leu Pro Pro His Asn Ser Pro His Arg Asp Leu Ala Pro Arg Gly
145 150 155 160Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Leu Pro Ala Thr Pro Glu
165 170 175Ser Arg Gly Ile Pro Ala Pro Glu Pro Pro Asn Val Gly Ser Ser Asp
180 185 190Pro Leu Ser Met Val Gly Pro Leu Gln Gly Gln Ser Pro Ser Tyr Thr
195 200 205

US 9,475,856 B2

265

266

-continued

Ser

```

<210> SEQ ID NO 127
<211> LENGTH: 198
<212> TYPE: PRT
<213> ORGANISM: Takifugu rubripes

<400> SEQUENCE: 127

Phe Ile Tyr Leu Phe Ile Gln Thr Ala Leu Phe Ser Pro Ser Lys Trp
1 5 10 15

Phe Asn Phe Tyr Leu Pro Asp Ser Asn Pro Leu Leu Ser Phe Asp Ser
20 25 30

His Gly Arg Gly Ile His Leu Tyr Thr Asp Asn Gln Arg Arg Gly Met
35 40 45

Tyr Leu Gln Met Ser Thr Asp Gly Ser Val Ser Gly Ser Asp Val Gln
50 55 60

Thr Ala Asn Ser Val Leu Glu Leu Lys Ser Val Arg Asn Gly His Val
65 70 75 80

Val Ile Arg Gly Lys Ser Ser Ser Leu Phe Leu Cys Met Asp Ser Arg
85 90 95

Gly Arg Leu Trp Gly Gln Arg His Pro Thr Glu Ala Asp Cys Thr Phe
100 105 110

Arg Glu Val Leu Leu Ala Asp Gly Tyr Thr Arg Phe Leu Ser Leu His
115 120 125

Asn Gly Thr Pro Val Ser Leu Ala Pro Lys Gln Ser Pro Asp Gln His
130 135 140

Thr Val Pro Phe Thr Arg Phe Leu Pro Leu Arg Asn Thr Leu Ala Glu
145 150 155 160

Glu Ser Met Ser Glu Pro Pro Ser Asn Gln Gln Arg Tyr Phe Asn Ile
165 170 175

Asp Ser Asp Asp Leu Leu Gly Met Asp Leu Asn Ala Met Val Ser Pro
180 185 190

Gln Phe Ser Gly Asp Lys
195

```

```

<210> SEQ ID NO 128
<211> LENGTH: 203
<212> TYPE: PRT
<213> ORGANISM: Dipodomys ordii

<400> SEQUENCE: 128

Met Asp Gln Ala Lys Thr Arg Val Gly Ala Arg Gly Leu Gly Gly Leu
1           5           10          15

Val Leu Ala Val Ile Ile Leu Gly Ala Cys Lys Ala Arg Pro Ile Pro
20          25          30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Leu Arg His
35          40          45

Leu Tyr Thr Asp Asp Thr Gln Glu Thr Glu Ala His Leu Glu Ile Arg
50          55          60

Ala Asp Gly Thr Val Val Gly Thr Ala His Arg Ser Pro Glu Ser Leu
65          70          75          80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Ile
85          90          95

Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Thr Leu Tyr Gly
100         105         110

```

-continued

Ser Leu His Phe Asp Pro Glu Val Cys Ser Phe Gln Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Ile Tyr Arg Ser Glu Ala Leu Gly Leu Pro Leu
130 135 140

Arg Leu Ser Pro Asp Pro Ala Pro Trp Gly Pro Ala Arg Phe Leu Pro
145 150 155 160

Leu Pro Gly Val Pro Pro Ala Pro Pro Glu Pro Pro Gly Ile Leu Ala
165 170 175

Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly
180 185 190

Leu Leu Gln Gly Arg Ser Pro Ser Tyr Ala Ser
195 200

<210> SEQ ID NO 129

<211> LENGTH: 184

<212> TYPE: PRT

<213> ORGANISM: Echinops telfairi

<400> SEQUENCE: 129

Met Gly Cys Thr Lys Ser Gly Trp Lys Ser Pro Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Ser Leu Leu Leu Gly Gly Cys Gly Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Thr Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Ala Asp Gly Thr Val Gly Gly Val Ala His Gln Ser Pro Glu Lys Phe
65 70 75 80

Leu Ser Gln Trp Arg Glu Lys Pro Leu Arg Ser Leu His Phe Asp Pro
85 90 95

Ala Ala Cys Ser Phe Arg Glu Lys Leu Leu Glu Asp Gly Tyr Asn Leu
100 105 110

Tyr His Ser Glu Thr His Gly Leu Pro Leu Arg Leu Pro Pro Arg Gly
115 120 125

Gly Asp Pro Ser Ser Gln Pro Gly Ala Arg Phe Pro Pro Leu Pro Gly
130 135 140

Gln Leu Pro Gln Leu Gln Glu Thr Pro Gly Val Leu Ala Pro Glu Pro
145 150 155 160

Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Trp Arg
165 170 175

Gly Gln Ser Pro Ser Tyr Ala Ser
180

<210> SEQ ID NO 130

<211> LENGTH: 210

<212> TYPE: PRT

<213> ORGANISM: Macaca mulatta

<400> SEQUENCE: 130

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Pro
1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

-continued

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
 50 55 60
 Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu Ser Glu
 65 70 75 80
 Cys Gly Pro Glu Pro Gly Ser Glu Gly Gly Ala Val Gly Gly Ala
 85 90 95
 Glu Gly Pro Gly Leu Leu Gly Leu Arg Glu Ala Gly Leu Gly Pro Gly
 100 105 110
 Ser Trp Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu
 115 120 125
 Leu Glu Asn Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro
 130 135 140
 Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Ser Gln
 145 150 155 160
 Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Pro Pro
 165 170 175
 Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser
 180 185 190
 Asp Pro Leu Ser Met Val Gly Pro Ser Gln Ala Arg Ser Pro Ser Tyr
 195 200 205
 Ala Ser
 210

<210> SEQ ID NO 131
 <211> LENGTH: 193
 <212> TYPE: PRT
 <213> ORGANISM: Microcebus murinus

<400> SEQUENCE: 131

Met Gly Trp Asp Glu Ala Gly Ala Gly Phe Glu His Pro Gly Leu Trp
 1 5 10 15
 Phe Pro Met Leu Gly Val Leu Leu Gly Ala Cys Gln Ala Tyr Pro
 20 25 30
 Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln
 35 40 45
 Arg His Leu Tyr Thr Asp Asp Ile Gln Glu Thr Glu Ala His Leu Glu
 50 55 60
 Ile Arg Ala Asp Gly Thr Val Val Gly Ala Ala Arg Gln Ser Pro Glu
 65 70 75 80
 Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95
 Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
 100 105 110
 Ser Leu His Phe Asp Pro Glu Cys Ser Phe Arg Glu Leu Leu Leu Glu
 115 120 125
 Asp Gly Tyr Asn Val Tyr Cys Pro Tyr Leu Pro Leu His Leu Ser Pro
 130 135 140
 Arg Ile Glu Leu Ala Gly Ser Arg Ser Ala Leu Pro Leu Pro Pro Ala
 145 150 155 160
 Pro Glu Arg Arg Ile Leu Ala Pro Glu Pro Pro Asp Gly Ser Ser Asp
 165 170 175
 Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
 180 185 190

Ser

-continued

<210> SEQ ID NO 132
<211> LENGTH: 204
<212> TYPE: PRT
<213> ORGANISM: Ochotona princeps

<400> SEQUENCE: 132

Lys	Asp	Met	Asp	Gly	Leu	Gln	Pro	Pro	Gly	Leu	Arg	Val	Pro	Val	Leu
1				5			10				15				
Ala	Ala	Leu	Leu	Leu	Gly	Val	Gly	Gln	Ala	Arg	Pro	Ile	Pro	Asp	Ser
		20				25				30					
Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	His	Leu	Tyr
		35				40			45						
Thr	Asp	Asp	Ala	Gln	Glu	Ser	Glu	Val	His	Leu	Glu	Ile	Arg	Ala	Asp
		50				55			60						
Gly	Thr	Val	Ala	Gly	Thr	Ala	Arg	Arg	Ser	Pro	Glu	Ser	Leu	Leu	Glu
		65				70			75				80		
Met	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	Ile	Leu	Gly	Val	His	Thr
		85					90				95				
Ser	Arg	Phe	Leu	Cys	Gln	Arg	Pro	Asp	Gly	Thr	Leu	Tyr	Gly	Ser	Leu
		100					105			110					
His	Phe	Asp	His	Lys	Ala	Cys	Ser	Phe	Arg	Glu	Gln	Leu	Leu	Glu	Asp
		115					120			125					
Gly	Tyr	Asn	Val	Tyr	His	Ser	Glu	Thr	His	Gly	Leu	Pro	Leu	Arg	Leu
		130				135			140						
Ser	Pro	Asp	Arg	Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Leu	Pro
		145				150			155			160			
Gly	Pro	Pro	Pro	Asp	Leu	Leu	Val	Pro	Pro	Leu	Pro	Pro	Asp	Val	Leu
		165					170			175					
Ala	Pro	Glu	Pro	Pro	Asp	Val	Asp	Ser	Pro	Asp	Pro	Leu	Ser	Met	Val
		180					185			190					
Gly	Pro	Leu	Gln	Gly	Gln	Ser	Pro	Ser	Tyr	Thr	Ser				
		195						200							

<210> SEQ ID NO 133
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Xiphophorus maculatus

<400> SEQUENCE: 133

Cys	Pro	Phe	Pro	Phe	Leu	Phe	Leu	Ile	Leu	Ser	Leu	Pro	Phe	Phe	Ser
1					5			10			15				
Ser	Ser	Phe	Tyr	Ile	Pro	Glu	Ser	Asn	Pro	Ile	Phe	Ala	Phe	Arg	Asn
		20				25			30						
Gln	Leu	Arg	Glu	Val	His	Leu	Tyr	Thr	Glu	Asn	His	Arg	Arg	Gly	Leu
		35				40			45						
Tyr	Val	Glu	Ile	His	Leu	Asp	Gly	Arg	Val	Thr	Gly	Ser	Asp	Ala	Gln
		50				55			60						
Ser	Pro	Tyr	Ser	Val	Leu	Gln	Ile	Lys	Ser	Val	Lys	Pro	Gly	His	Val
		65				70			75			80			
Val	Ile	Lys	Gly	Gln	Thr	Ser	Ser	Leu	Phe	Leu	Cys	Met	Asp	Asp	Ser
		85					90			95					
Gly	Asn	Leu	Arg	Gly	Gln	Thr	Thr	Tyr	Asp	Glu	Ala	Asp	Cys	Ser	Phe
		100					105			110					
Arg	Glu	Leu	Leu	Leu	Ala	Asp	Gly	Tyr	Thr	Arg	Phe	Leu	Asn	Ser	Gln
		115					120			125					

-continued

His Gly Val Pro Leu Ser Leu Ala Ser Arg Asn Ser Pro Asp Arg His
130 135 140

Ser Val Pro Phe Thr Arg Phe Leu Pro Leu Arg Asn Thr Leu Thr Val
145 150 155 160

Ser Glu Glu Ser Thr Lys Thr Gln Arg Asp Phe Asn Leu Asp Ser Asp
165 170 175

Asp Leu Leu Gly Met Gly
180

<210> SEQ ID NO 134

<211> LENGTH: 192

<212> TYPE: PRT

<213> ORGANISM: Gasterosteus aculeatus

<400> SEQUENCE: 134

Ser Leu Leu Leu Met Val Pro Leu Pro Phe Cys Ser Ser Phe Tyr Leu
1 5 10 15

Thr Asp Ser Ser Pro Leu Leu Pro Phe Asn Asn Gln Val Lys Glu Val
20 25 30

His Leu Tyr Thr Ala Glu Asn His Arg Arg Ala Met Tyr Leu Gln Ile
35 40 45

Ala Leu Asp Gly Ser Val Ser Gly Ser Asp Ala Arg Ser Thr Tyr Ser
50 55 60

Val Leu Gln Leu Lys Ser Ile Gln Pro Gly His Val Val Ile Arg Gly
65 70 75 80

Lys Ala Ser Ser Met Phe Leu Cys Val Asp Ser Gly Gly Arg Leu Arg
85 90 95

Gly Gln Gly Pro Tyr Ser Glu Ala Asp Cys Ser Phe Arg Glu Leu Leu
100 105 110

Leu Gly Asp Gly Tyr Thr Arg Phe Leu Ser Ser Gln His Gly Ser Pro
115 120 125

Leu Ser Leu Ala Ser Arg Pro Ser Pro Asp Pro Asn Ser Val Pro Phe
130 135 140

Thr Arg Phe Leu Pro Ile Arg Thr Ala Pro Glu Ala Glu Ser Val Ile
145 150 155 160

Glu Glu Pro Pro Ser Asn Gln Arg Tyr Val Asn Val Asp Ser Glu Asp
165 170 175

Leu Leu Gly Met Gly Leu Asn Thr Val Val Ser Pro Gln Phe Ser Ala
180 185 190

<210> SEQ ID NO 135

<211> LENGTH: 165

<212> TYPE: PRT

<213> ORGANISM: Sarcophilus harrisii

<400> SEQUENCE: 135

Val Ser Ala Met Gly Leu Arg Glu Arg Ala Pro Arg Tyr Leu Ala Pro
1 5 10 15

Leu Leu Ser Leu Leu Ala Cys Arg Ala Ser Gly His Pro Leu Pro
20 25 30

Asp Ser Ser Pro Met Leu Leu Phe Gly Gly Gln Val Arg Leu Arg His
35 40 45

Leu Tyr Thr Asp Val Gly Gln Glu Ala Glu Ala His Val Glu Leu Ala
50 55 60

Ser Asp Gly Thr Val Arg Ala Ala Ala Arg Arg Ser Pro Asn Ser Leu
65 70 75 80

-continued

Leu Glu Leu Lys Ala Val Lys Pro Gly Ile Val Arg Ile Leu Ala Val
85 90 95

His Ser Ser Arg Phe Leu Cys Met Arg Pro Asn Gly Glu Leu Tyr Gly
100 105 110

Ala Ile His Tyr Asp Pro Ser Ala Cys Asn Phe Arg Glu Arg Leu Leu
115 120 125

Gly Asp Gly Tyr Asn Val Tyr Glu Ser Glu Ala His Gly Arg Thr Leu
130 135 140

Arg Leu Pro Pro Lys Ala Ala Pro Gly Pro Ala Gly Pro Ser Arg Phe
145 150 155 160

Leu Pro Leu Pro Gly
165

<210> SEQ_ID NO 136

<211> LENGTH: 208

<212> TYPE: PRT

<213> ORGANISM: Macropus eugenii

<400> SEQUENCE: 136

Thr Glu Glu Pro Ser Thr Gly Ser Arg His Leu Gly Gln Trp Ala Pro
1 5 10 15

Gly Leu Pro Gly Pro Leu Leu Ser Leu Leu Leu Ala Tyr Arg Gly Trp
20 25 30

Gly Ser Pro Ile Pro Asp Ser Ser Pro Met Leu Leu Phe Gly Gly Gln
35 40 45

Val Arg Leu Arg His Leu Tyr Thr Asp Asp Gly Gln Asp Thr Glu Ala
50 55 60

His Val Glu Leu Gly Pro Asp Gly Val Val Arg Ala Val Ala Glu Arg
65 70 75 80

Ser Pro Asn Ser Leu Leu Glu Leu Lys Ala Val Lys Pro Gly Val Ile
85 90 95

Arg Ile Leu Ala Val Gln Ser Ser Arg Phe Leu Cys Met Arg Pro Asn
100 105 110

Gly Glu Leu Tyr Gly Ala Val His Tyr Asp Pro Ser Ala Cys Asn Phe
115 120 125

Arg Glu His Leu Leu Gly Asp Gly Tyr Asn Val Tyr Glu Ser Glu Thr
130 135 140

His Arg Arg Thr Leu Arg Leu Ser Pro Ser Leu Gly Gln Ala Gly Pro
145 150 155 160

Ser Arg Phe Leu Pro Leu Pro Gly Asp Trp Leu Pro Gly Pro Asp Pro
165 170 175

Pro Trp Ala Gln Gly Pro Glu Pro Pro Asp Val Gly Ser Ala Asp Pro
180 185 190

Leu Ser Met Val Gly Ala Val Gln Gly Leu Ser Pro Ser Tyr Ser Ser
195 200 205

<210> SEQ_ID NO 137

<211> LENGTH: 160

<212> TYPE: PRT

<213> ORGANISM: Xenopus tropicalis

<400> SEQUENCE: 137

Arg Gly Gly Arg Thr Lys Lys Thr Leu Leu Arg Lys Trp Leu Cys
1 5 10 15

Leu Leu Ala Ile Met Leu Ser Arg Ser Arg Phe Ser Leu Ala Asn Pro
20 25 30

-continued

Ile Gln Asn Ser Asn Pro Ile Leu Ser Asn Asp Asn Gln Val Arg Thr
35 40 45

Gln Tyr Leu Tyr Thr Asp Asn Asn Met His Leu Tyr Leu Gln Ile
50 55 60

Thr His Asn Gly Val Val Thr Gly Thr Glu Glu Lys Asn Asp Tyr Gly
65 70 75 80

Val Leu Glu Ile Lys Ala Val Lys Ala Gly Val Val Val Ile Lys Gly
85 90 95

Ile Arg Ser Asn Leu Tyr Leu Cys Met Asp Ser Arg His Gln Leu Tyr
100 105 110

Ala Ser Ala Tyr Asp Lys Asp Asp Cys His Phe His Glu Lys Ile Thr
115 120 125

Pro Asp Asn Tyr Asn Met Tyr Ser Ser Glu Lys His Ser Glu Tyr Val
130 135 140

Ser Leu Ala Pro Leu Lys Gly Ser Gln Met Ala Arg Phe Leu Pro Ile
145 150 155 160

<210> SEQ ID NO 138

<211> LENGTH: 194

<212> TYPE: PRT

<213> ORGANISM: Danio rerio

<400> SEQUENCE: 138

Met Leu Leu Ala Cys Phe Phe Ile Phe Phe Ala Leu Phe Pro His Leu
1 5 10 15

Arg Trp Cys Met Tyr Val Pro Ala Gln Asn Val Leu Leu Gln Phe Gly
20 25 30

Thr Gln Val Arg Glu Arg Leu Leu Tyr Thr Asp Gly Leu Phe Leu Glu
35 40 45

Met Asn Pro Asp Gly Ser Val Lys Gly Ser Pro Glu Lys Asn Leu Asn
50 55 60

Cys Val Leu Glu Leu Arg Ser Val Lys Ala Gly Glu Thr Val Ile Gln
65 70 75 80

Ser Ala Ala Thr Ser Leu Tyr Leu Cys Val Asp Asp Gln Asp Lys Leu
85 90 95

Lys Gly Gln His His Tyr Ser Ala Leu Asp Cys Thr Phe Gln Glu Leu
100 105 110

Leu Leu Asp Gly Tyr Ser Phe Phe Leu Ser Pro His Thr Asn Leu Pro
115 120 125

Val Ser Leu Leu Ser Lys Arg Gln Lys His Gly Asn Pro Leu Ser Arg
130 135 140

Phe Leu Pro Val Ser Arg Ala Glu Asp Ser Arg Thr Gln Glu Val Lys
145 150 155 160

Gln Tyr Ile Gln Asp Ile Asn Leu Asp Ser Asp Asp Pro Leu Gly Met
165 170 175

Gly His Arg Ser His Leu Gln Thr Val Phe Ser Pro Ser Leu His Thr
180 185 190

Lys Lys

<210> SEQ ID NO 139

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Bos grunniens mutus

<400> SEQUENCE: 139

Met Gly Trp Asp Glu Ala Lys Phe Lys His Leu Gly Leu Trp Val Pro

US 9,475,856 B2

279**280**

-continued

1	5	10	15
Val Leu Ala Val Leu Leu Leu Gly Thr Cys Arg Ala His Pro Ile Pro			
20	25	30	
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr			
35	40	45	
Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg			
50	55	60	
Ala Asp Gly Thr Val Val Gly Ala Ala Arg Gln Ser Pro Glu Ser Leu			
65	70	75	80
Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val			
85	90	95	
Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Lys Leu Tyr Gly			
100	105	110	
Ser Leu His Phe Asp Pro Lys Ala Cys Ser Phe Arg Glu Leu Leu Leu			
115	120	125	
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Thr Leu Gly Leu Pro Leu			
130	135	140	
Arg Leu Pro Pro Gln Arg Ser Ser Asn Arg Asp Pro Ala Pro Arg Gly			
145	150	155	160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Ala Glu Pro Pro Asp			
165	170	175	
Pro Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp			
180	185	190	
Pro Leu Ser Met Val Gly Pro Ser Tyr Gly Arg Ser Pro Ser Tyr Thr			
195	200	205	

Ser

<210> SEQ ID NO 140
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Saimiri boliviensis boliviensis

<400> SEQUENCE: 140

Met Gly Ser Glu Glu Val Ala Leu Glu Arg Pro Ala Leu Trp Val Ser			
1	5	10	15
Val Leu Ala Gly Leu Leu Leu Gly Thr Cys Gln Ala Tyr Pro Ile Pro			
20	25	30	
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr			
35	40	45	
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg			
50	55	60	
Glu Asp Gly Thr Val Ala Gly Ala Ala His Gln Ser Pro Glu Ser Leu			
65	70	75	80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val			
85	90	95	
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly			
100	105	110	
Ser Leu Tyr Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu			
115	120	125	
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Val Ala His Ser Leu Pro Leu			
130	135	140	
His Leu Pro Gly Gly Arg Ser Pro Pro Trp Asp Pro Ala Pro Arg Gly			
145	150	155	160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Glu Pro Pro Glu			

-continued

165	170	175
-----	-----	-----

Ala Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Gln Ser Pro Ser Tyr Thr
195 200 205

Ser

```
<210> SEQ ID NO 141
<211> LENGTH: 210
<212> TYPE: PRT
<213> ORGANISM: Callithrix jacchus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (64)..(64)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid

<400> SEQUENCE: 141
```

Met Gly Ser Glu Glu Val Gly Leu Glu His Pro Ala Leu Trp Val Ser
1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Thr Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Lys Glu Ala His Leu Glu Ile Xaa
50 55 60

Glu Asp Gly Thr Val Ala Gly Ala Ala Thr Lys Val Pro Lys Val Ser
65 70 75 80

Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly
85 90 95

Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr
100 105 110

Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu
115 120 125

Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Val Ala His Gly Leu Pro
130 135 140

Leu His Leu Pro Glu Ser Arg Ser Pro Pro Arg Asp Pro Ala Pro Arg
145 150 155 160

Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Glu Pro Pro
165 170 175

Glu Pro Pro Gly Ile Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser
180 185 190

Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Gln Ser Pro Ser Tyr
195 200 205

Ala Ser
210

```
<210> SEQ ID NO 142
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Tupaia chinensis
```

<400> SEQUENCE: 142

Met Gly Trp Asp Lys Ala Arg Phe Glu His Leu Gly Ala Trp Ala Pro
1 5 10 15

Val Leu Ala Val Leu Leu Leu Gly Ala Cys Gln Ala Tyr Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr

-continued

35	40	45
Leu Tyr Thr Asp Asp Thr Gln Asp Thr Glu Ala His Leu Glu Ile Arg		
50	55	60
Ala Asp Gly Thr Val Val Gly Ala Ala His Gln Ser Pro Glu Ser Leu		
65	70	75
80		
Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val		
85	90	95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly		
100	105	110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu		
115	120	125
Glu Asp Gly Tyr Asn Ile Tyr Gln Ser Glu Ala Arg Gly Leu Pro Leu		
130	135	140
Arg Leu Pro Pro His Asp Ser Pro His Arg Asp Arg Thr Pro Gln Gly		
145	150	155
160		
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Leu Val Pro Pro Glu		
165	170	175
Leu Pro Gly Val Leu Ala Leu Glu Pro Pro Asp Val Gly Ser Ser Asp		
180	185	190
Pro Leu Ser Met Met Gly Pro Ser Gln Gly Gln Ser Pro Ser Tyr Ala		
195	200	205

Ser

<210> SEQ_ID NO 143
<211> LENGTH: 188
<212> TYPE: PRT
<213> ORGANISM: *Papio anubis*

<400> SEQUENCE: 143

Met	Asp	Ser	Asp	Glu	Thr	Gly	Phe	Glu	His	Ser	Gly	Leu	Trp	Val	Pro
1				5					10					15	
Val	Leu	Ala	Gly	Leu	Leu	Leu	Gly	Ala	Cys	Gln	Ala	His	Pro	Ile	Pro
				20					25					30	
Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	Tyr
						35			40			45			
Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His	Leu	Glu	Ile	Arg
						50			55			60			
Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	His	Gln	Ser	Pro	Glu	Ser	Lys
					65			70			75			80	
Cys	Gly	Pro	Glu	Pro	Gly	Ser	Glu	Gly	Gly	Gly	Ala	Leu	His	Phe	Asp
						85				90				95	
Pro	Glu	Ala	Cys	Ser	Phe	Arg	Glu	Leu	Leu	Leu	Glu	Asn	Gly	Tyr	Asn
					100				105				110		
Val	Tyr	Gln	Ser	Glu	Ala	His	Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn
					115				120			125			
Lys	Ser	Pro	His	Arg	Asp	Pro	Ala	Ser	Arg	Gly	Pro	Ala	Arg	Phe	Leu
						130			135			140			
Pro	Leu	Pro	Gly	Leu	Pro	Pro	Ala	Pro	Pro	Glu	Pro	Pro	Gly	Ile	Leu
					145				150			155			160
Ala	Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val
						165				170				175	
Gly	Pro	Ser	Gln	Ala	Arg	Ser	Pro	Ser	Tyr	Ala	Ser				
						180			185						

-continued

<210> SEQ_ID NO 144
 <211> LENGTH: 208
 <212> TYPE: PRT
 <213> ORGANISM: Pteropus alecto
 <400> SEQUENCE: 144

Met	Gly	Trp	Gly	Lys	Ala	Arg	Leu	Gln	His	Pro	Gly	Leu	Trp	Gly	Pro
1				5			10						15		

Val Leu Ala Val Leu Leu Gly Ala Cys Gln Ala His Pro Ile Leu Asp
 20 25 30

Ser Ser Pro Leu Phe Gln Phe Gly Ser Gln Val Arg Arg Arg Tyr Leu
 35 40 45

Tyr Thr Asp Asp Ala Gln Asp Thr Glu Ala His Leu Glu Ile Arg Ala
 50 55 60

Asp Gly Thr Val Ala Gly Ala Ala Arg Arg Ser Pro Glu Ser Leu Leu
 65 70 75 80

Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Val Leu Gly Val Lys
 85 90 95

Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Thr Leu Tyr Gly Ser
 100 105 110

Leu His Phe Asp Pro Ala Ala Cys Ser Phe Arg Glu Leu Leu Lys
 115 120 125

Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala Leu Ala Arg Pro Leu Arg
 130 135 140

Leu Pro Pro Tyr Ser Ser Pro Ser Ser Asp Pro Ala Arg Arg Gly Pro
 145 150 155 160

Ala Arg Phe Leu Pro Leu Pro Gly Pro Pro Pro Glu Pro Pro Gln Pro
 165 170 175

Pro Gly Arg Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp Pro
 180 185 190

Leu Ser Met Val Trp Pro Ser Arg Gly Arg Ser Pro Ser Tyr Thr Ser
 195 200 205

<210> SEQ_ID NO 145
 <211> LENGTH: 210
 <212> TYPE: PRT
 <213> ORGANISM: Heterocephalus glaber
 <400> SEQUENCE: 145

Met	Asp	Trp	Ala	Arg	Ala	Glu	Ser	Glu	Arg	Pro	Gly	Leu	Trp	Val	Pro
1				5			10					15			

Ala Val Leu Ala Val Leu Leu Gly Ala Cys Gln Ala His Pro Ile
 20 25 30

Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg
 35 40 45

His Leu Tyr Thr Asp Asp Ala Gln Asp Thr Glu Val His Leu Glu Ile
 50 55 60

Arg Ala Asp Gly Ser Val Gly Gly Ala Ala His Arg Ser Pro Glu Ser
 65 70 75 80

Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly
 85 90 95

Val Arg Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Thr Leu Tyr
 100 105 110

Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu
 115 120 125

Leu Ala Asp Gly Tyr Asn Ile Tyr Gln Ser Glu Ala Tyr Gly Leu Pro

-continued

130	135	140
Leu Arg Met Leu Pro Ser Asp Ser Ala Ser Arg Asp Pro Val Pro Pro		
145	150	155
Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu His Pro Pro Pro Leu		
165	170	175
Glu Pro Pro Gly Met Leu Pro Pro Glu Pro Pro Asp Val Gly Ser Ser		
180	185	190
Asp Pro Leu Ser Met Val Gly Pro Leu Gln Gly Arg Ser Pro Ser Tyr		
195	200	205
Ala Phe		
210		

<210> SEQ ID NO 146
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Cricetulus griseus

<400> SEQUENCE: 146

Met Asp Trp Met Lys Ser Gly Val Gly Val Pro Gly Leu Trp Val Pro			
1	5	10	15
Leu Leu Pro Ile Phe Leu Leu Gly Val Ser Gln Ala His Pro Ile Pro			
20	25	30	
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg His Arg His			
35	40	45	
Leu Tyr Thr Asp Asp Asn Gln Glu Thr Glu Val His Leu Glu Ile Arg			
50	55	60	
Gln Asp Gly Thr Val Ile Gly Thr Thr His Arg Ser Pro Glu Ser Leu			
65	70	75	80
Leu Glu Leu Lys Ala Leu Lys Pro Glu Val Ile Pro Val Leu Gly Val			
85	90	95	
Lys Ala Ser Arg Phe Leu Cys Gln Gln Pro Asp Gly Thr Leu Tyr Gly			
100	105	110	
Ser Pro His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu			
115	120	125	
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Val His Gly Leu Pro Leu			
130	135	140	

Arg Leu Pro Gln Arg Asp Ser Pro Asn Gln Ala Pro Ala Ser Trp Gly			
145	150	155	160
Pro Val Pro Pro Leu Pro Val Pro Gly Leu Leu His Gln Pro Gln Glu			
165	170	175	
Leu Pro Gly Phe Leu Ala Pro Glu Pro Pro Asp Val Gly Ser Ser Asp			
180	185	190	
Pro Leu Ser Met Val Gly Pro Leu Gln Gly Arg Ser Pro Ser Tyr Ala			
195	200	205	

Ser

<210> SEQ ID NO 147
<211> LENGTH: 192
<212> TYPE: PRT
<213> ORGANISM: Ovis aries

<400> SEQUENCE: 147

Met Gly Trp Asp Glu Ala Lys Phe Lys His Leu Gly Leu Trp Val Pro			
1	5	10	15
Val Leu Ala Val Leu Leu Gly Thr Cys Arg Ala His Pro Ile Pro			
20	25	30	

US 9,475,856 B2

289**290**

-continued

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
 35 40 45

Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
 50 55 60

Ala Asp Gly Thr Val Val Gly Ala Ala Arg Gln Ser Pro Glu Ser Leu
 65 70 75 80

Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Phe Gly Val
 85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Gly Pro Asp Gly Lys Leu Tyr Gly
 100 105 110

Ser Leu His Phe Asp Pro Lys Ala Cys Ser Phe Arg Glu Leu Leu Leu
 115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Thr Leu Gly Leu Pro Leu
 130 135 140

Arg Leu Pro Pro Gln Arg Ser Ser Asn Arg Asp Pro Ala Pro Arg Gly
 145 150 155 160

Pro Pro Lys Pro Gln Leu His Phe Leu Lys Thr Ser Ala Val Gln Tyr
 165 170 175

Trp Pro Arg Tyr Glu Lys Val Pro Ala Phe Leu His Pro Phe Pro Gly
 180 185 190

<210> SEQ ID NO 148

<211> LENGTH: 135

<212> TYPE: PRT

<213> ORGANISM: Pan paniscus

<400> SEQUENCE: 148

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
 1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
 20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
 35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
 50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
 65 70 75 80

Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
 100 105 110

Ser Val Ser Phe Gln Asp Pro Pro His His Pro Pro Cys Ser Ser Tyr
 115 120 125

Met Ser Pro Ser Gln Pro Gly
 130 135

<210> SEQ ID NO 149

<211> LENGTH: 116

<212> TYPE: PRT

<213> ORGANISM: Macaca fascicularis

<400> SEQUENCE: 149

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Pro
 1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
 20 25 30

-continued

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
 35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
 50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu Ser Leu
 65 70 75 80

Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
 85 90 95

Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu Tyr Gly
 100 105 110

Ser Val Ser Phe
 115

<210> SEQ ID NO 150
 <211> LENGTH: 104
 <212> TYPE: PRT
 <213> ORGANISM: Mesocricetus auratus

<400> SEQUENCE: 150

Val Ile Gln Ile Leu Gly Val Lys Ala Ala Arg Phe Pro Cys Gln Gln
 1 5 10 15

Pro Asp Gly Ser Leu Tyr Gly Ser Pro His Phe Asp Pro Glu Ala Cys
 20 25 30

Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
 35 40 45

Glu Ala His Gly Leu Pro Leu Arg Leu Pro Gln Arg Asp Ala Pro Ser
 50 55 60

Gln Pro Pro Ala Ser Trp Gly Pro Val Arg Phe Leu Pro Val Pro Gly
 65 70 75 80

Leu Phe Gln Pro Pro His Asp Leu Pro Gly Arg Pro Ala Pro Glu Pro
 85 90 95

Pro Asp Val Gly Ser Ser Asp Pro
 100

<210> SEQ ID NO 151
 <211> LENGTH: 151
 <212> TYPE: PRT
 <213> ORGANISM: Oreochromis niloticus

<400> SEQUENCE: 151

Met Tyr Leu Gln Met Asn Met Asp Gly Arg Val Thr Gly Ser Asp Ala
 1 5 10 15

Gln Thr Pro Tyr Ser Leu Met Gln Leu Lys Ser Val Lys Pro Gly His
 20 25 30

Val Ile Ile Lys Gly Pro Ser Ser Ser Leu Phe Leu Cys Val Asp Ser
 35 40 45

Glu Gly Asn Leu Arg Gly Gln Ser His Tyr Ser Glu Thr Ser Cys Thr
 50 55 60

Phe Arg Glu Met Leu Leu Ala Asp Gly Tyr Thr Arg Phe Ile Ser Ser
 65 70 75 80

Gln Tyr Gly Phe Pro Met Ser Leu Ala Ser Arg His Ser Pro Asp Arg
 85 90 95

His Ala Leu Pro Phe Thr Arg Phe Leu Pro Leu Arg Asn Asn Leu Lys
 100 105 110

Thr Asp Ser Val Ser Glu Gln Leu Pro Asn Asn Gln Arg Leu Phe Asn
 115 120 125

-continued

Val Asp Ser Asp Asp Leu Leu Gly Met Gly Leu Asn Ser Met Gly Ser
130 135 140

Pro Gln Phe Ser Met Asp Lys
145 150

<210> SEQ ID NO 152

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 152

Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
1 5 10 15

Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
20 25 30

Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
35 40 45

Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
50 55 60

Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65 70 75 80

Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
85 90 95

Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly Ala Leu Tyr Gly
100 105 110

Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
115 120 125

Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
130 135 140

His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
165 170 175

Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205

Ser

<210> SEQ ID NO 153

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 153

atggactcgg acgagacccgg gttcgagcac tcaggactgt gggtttctgt gctggcttgt 60

cttctgtctgg gagcctgcga ggcacacccc atccctgact ccagtctct cctgcaattc 120

gggggccaag tccggcagcg gtacctctac acagatgtat cccagcagac agaagccac 180

ctggagatca gggaggatgg gacggtggttgg accagagccc cgaaagtctc 240

ctgcagctga aagccttcaa gcccggagtt attcaaattct tgggagtcaa gacatccagg 300

tccctgtgcc agcggccaga tggggccctg tatggatcgcc tccacttga ccctgaggcc 360

tgcagcttcc gggagctgct tcttgaggac ggataacaatg ttaccatgc cgaagccac 420

ggcctcccgcc tgcacccgtcc agggaaacaag tccccacacc gggaccctgc accccgagga 480

-continued

ccagctcgct tcctgccact accaggcctg ccccccgcac tccggagcc acccggaatc	540
ctggcccccc agccccccga tgtgggctcc tggaccctc tgagcatggt gggaccttcc	600
cagggccgaa gccccagcta cgcttcctga	630

<210> SEQ ID NO 154
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Pongo abelii

<400> SEQUENCE: 154

atggactcg acgagacccgg gttcgagcac tcaggactgt gggttcctgt gctggcttgt	60
cttctgtgg gagcctgcca ggcacacccc atccctgact ccagtcctct cctgcaattc	120
gggggccaag tccggcagcg gtacctctac acagatgtg cccagcagac agaagccac	180
ctggagatca gggaggatgg gacgggtgggg ggctgtgtg accagagccc cgaaagtctc	240
ctgcagctga aagcctgaa gcccggagtt attcaaattct tgggagtcaa gacatccagg	300
tccctgtgcc agaggccaga tggggccctg tatggatcgc tccactttga ccctgaggcc	360
tgcagettcc gggagctgtct tcttgaggac ggataacaatg tttatcgtc cgaggcccat	420
ggcctcccgcc tgcacactgcc gggaaacaag tccccacacc gggaccctgc accccgagga	480
ccagctcgct tcctgccact accaggcctg ccccccgcac ccccaagagcc gcccgaaatc	540
ctggcccccc agccccccga tgtgggctcc tggaccctc tgagcatggt gggaccttcc	600
cagggccgaa gccccagcta tgcttcctga	630

<210> SEQ ID NO 155
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 155

atggactcg acgagacccgg gttcgagcac tcaggactgt gggttcctgt gctggcttgt	60
cttctgttag gagcctgcca ggcacacccc atccctgact ccagtcctct cctgcaattc	120
gggggccaag tccggcagcg gtacctctac acagatgtg cccagcagac agaagccac	180
ctggagatca gggaggatgg gacgggtgggg ggctgtgtg accagagccc cgaaagtctc	240
ctgcagctga aagcctgaa gcccggagtt attcaaattct tgggagtcaa gacatccagg	300
tccctgtgcc agaggccaga tggggccctg tatggatcgc tccactttga ccctgaggcc	360
tgcagettcc gggagctgtct tcttgaggac ggataacaatg tttaccagtc cgaggcccat	420
ggcctcccgcc tgcacactgcc gggaaacaag tccccacacc gggaccctgc accccgagga	480
ccagctcgct tcctgccact accaggcctg ccccccgcac ccccaagagcc acccgaaatc	540
ctggcccccc agccccccga tgtgggctcc tggaccctc tgagcatggt gggaccttcc	600
cagggccgaa gccccagcta cacttcctga	630

<210> SEQ ID NO 156
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Canis lupus familiaris

<400> SEQUENCE: 156

atgggctggg ccgaggccgg gttcgagcac ctgggactgt gggtccctgt gctggctgt	60
cttttgcgtgg aagcctgccc ggcacatccg atccctgact ccagccccct cctacaattt	120

-continued

ggaggtcaag ttcgacagcg gtacctctac accgacgtat cccaggagac agaggccac	180
ctagagatca gggccgatgg cacagtggtg ggggctgccc gccagagccc tgaaagtctc	240
ctggagatca aagccctaaa gccaggggtc attcaaattct tgggacttcaa aacatccagg	300
ttcctgtgcc agggcccgaga tgggacacta tatggctcg tccatttgcg ccctgtggcc	360
tgcagttcc gagaactgtcttgggat gggtacaaca tctaccactc cgagaccctt	420
ggtctccgcg ttcgectgcg cccccacaac tccgcatacc gggacttggc accccgcggg	480
cctgcccgtc tcctgccact gccaggcctg ctccagcac cccagagcc tccaggatc	540
ctggcccccgg agccctctga cgtgggctcc tgggaccctc tgagcatggt ggggcttca	600
caggggccgga gtcccagcta tgcttcctaa	630

<210> SEQ ID NO 157

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Bos taurus

<400> SEQUENCE: 157

atgggctggg acgaggccaa gttcaagcac ttgggactgt gggtccctgt gctggctgtc	60
ctcctgtctag gaacctgccc ggccatccc attccagact ccagccccc cctccagttt	120
gggggccaag tccgcccagcg gtacctctac acggatgtat cccaggagac agaggccac	180
ctggagatca gggccgatgg cacagtggtg ggggctggc gccagagccc cgaaagtctc	240
ttggagatca aagccctgaa gccaggcgctc attcagatct tgggactttaa aacatccagg	300
tttctgtgcc agggggccaga tgggaaagctg tacggatcgc tgcacttga ccccaagcc	360
tgcagtttc gggagatgtct tcttgaagat ggataacaacg tctaccagtc ggagaccctg	420
ggcctccac tccgcctgccc ccccaagcgcc tgcgttcaacc gggaccggc cccgcgggga	480
cctgctcgct tccttccact gccggggctg cccggggcgc cccggatcc tccaggatc	540
ttggcccccgg agccctccgaa cgtgggctcc tgggatcccc tgagatggt gggaccctcg	600
tatggccgaa gccccagcta cacttcttga	630

<210> SEQ ID NO 158

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Equus caballus

<400> SEQUENCE: 158

atggactggg acaagacggg gttcaagtac cagggactgt gggtccctgt gctggctgtc	60
cttctgtctgg gagcctgcca gtcacacccc atccctgact ccagtcctt cctccaattc	120
gggggccaag tcaggcagcg ccacccatctac acagatgtat cccaggagac agaggccac	180
ctggagatca gggctgacgg cactgtggca ggggctgtcc accggagccc agaaagtctc	240
ttggagatca aagccctgaa gccaggggta attcaaattct tgggacttcaa gacatccagg	300
tttctgtgcc agggggccaga cgggacgcgt tacggatcgc tccacttgcg cccctgtggcc	360
tgcagtttc gggagatgtct tcttgaagac ggataacaacg tttaccagtc tgagaccctt	420
ggcctccac tccgcctgccc ccaccacacg tcccccatacc agatccggc cccctgggca	480
cccgccccgt tcctgcccgtt gccaggcttt ccccaagcac ccccgagcc tccaggatc	540
ccggcccccgg agccccccgg a cgtgggctcc tgggacccttcc tgagatggt ggggcttca	600
cgcagccgga gccccagcta cacttccctga	630

US 9,475,856 B2

299**300**

-continued

<210> SEQ ID NO 159
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Ailuropoda melanoleuca

<400> SEQUENCE: 159

atgggctgg	acgaggccag	gtccgagcag	ctggggctgt	gggtccctgt	gctggctgtc	60
cttttgcgtgg	aagcttgcca	ggcacaccct	atcccctgact	ccagccccct	cctccaaattc	120
ggaggccaaag	ttcgacagcg	gtacccctac	acggacgtat	cccaggagac	agaggccac	180
ctagcgatca	gggctgtatgg	cacagtggtg	ggggctgcca	gccggagccc	agaaaagtctc	240
ttggagctga	aaggccctgaa	accgggggtc	attcaaattc	tgggagtgaa	aacatctagg	300
ttcctgtgcc	agggeccaga	tgggacactg	tacggatcgg	tccgcttcga	ccccgttagcc	360
tgcagettcc	gggaactgtct	cctggaggat	gggtacaaca	tctaccactc	tgagaccctc	420
ggcctccac	ttegectgcc	cgccccacaac	tctccatacc	gggactcggc	gccccggggg	480
cctgcccgt	tcttgcctt	gccaggcctg	cttccggtc	ccccggaccc	cccaggatc	540
ctggggccccg	agcctcccgaa	cgtggggctcc	tccggacccccc	tgagcatggt	ggggcctca	600
cagggccgaa	gtcccagcta	cgcttcctga				630

<210> SEQ ID NO 160
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Oryctolagus cuniculus

<400> SEQUENCE: 160

atggactggg	gcaaggccaa	gtgcggggcc	ccggggctgt	gggtccccgc	gctcgctgcc	60
ctgctgtgg	gggcctgcca	ggcacacccc	atccccact	ccagccccct	cctccagttt	120
ggggaccaag	tgccgcagca	gcacctgtac	acggacgtat	cgccaggaaac	agaagccac	180
ctggagatca	ggccggatgg	cacggatgg	ggggctgcc	ggaggagccc	agaaaagtctc	240
ttgcagatga	aaggcattaca	accggggatc	attcagatct	tgggggtc	gacgtccagg	300
ttcctctgcc	agaggccgga	tggcacgtc	tacggctc	tccacttcga	ccgcgaggcc	360
tgcagettcc	gggagctgtct	gcgtgaggat	gggtacaacg	tttacctctc	ggaggccctg	420
ggcctgcccc	tgcgectgtc	ccccggcagc	tcccccaegca	gggcgeccgc	ccccggggg	480
ccagcccgct	tcttgcctgt	gccccggctg	ccgcccagacc	tccggaaacc	gccaggctc	540
ctggccggcc	cgccccccgaa	tgtcgactcc	ccggacccccc	tgagcatggt	gcagectgcg	600
ctggaccaga	gccccagcta	cacccctcgt				630

<210> SEQ ID NO 161
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Gorilla gorilla

<400> SEQUENCE: 161

atggactcgg	acgagacccgg	gttcgagcac	tcaggactgt	gggtttctgt	gctggctgt	60
cttctgtgg	gagcctgcca	ggcacacccc	atcccctgact	ccagtcctct	cctgcattc	120
gggggccaag	tccggcagcg	gtacccctac	acagatgtat	cccaggacac	agaagccac	180
ctggagatca	gggaggatgg	gacggatgg	ggtgctgtct	accagagccc	tggaaagtctc	240
ctgcagctga	aaggcattgaa	gccgggagtt	attcaaattc	tgggagtgaa	gacatccagg	300
ttcctgtgcc	agaggccaga	tggggccctg	tatggatcgc	tccactttga	ccctgaggcc	360

-continued

tgcagcttcc	gggagctgct	tcttgaggac	ggataacaatg	tttaccagtc	cgaggcccac	420
ggcctcccg	tcacactg	ccccacacc	gggaccctgc	accccgagga		480
ccagctcg	tcctgcact	accaggc	ccccggcac	acccggaatc		540
ctggccccc	agcccccg	tgtggctcc	tggaccctc	tgagcatggt	gggaccttcc	600
cagggccgaa	gccccagcta	cgcttcctga				630

<210> SEQ ID NO 162

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Nomascus leucogenys

<400> SEQUENCE: 162

atggactcg	acgagac	gggtcgagcac	tcaggactgt	gggttcctgt	gctggcttgt	60
cttctgtgg	gagcctg	ggcacacccc	atccctgact	ccagtcctct	cctgcaattc	120
gggggccaag	tccggc	agcg	gtacctctac	acagatgtat	cccagcagac	180
ctggagatca	gggaggatgg	gacgggtgggg	ggcgctgtctg	accagagccc	tgaaagtctc	240
ctgcagctga	aagc	cttgc	aa	gcccggagtt	attcaa	300
ttcctatgc	agaggccaga	tggggccctg	tatggatcgc	tccacttga	ccctgaggcc	360
tgcagcttcc	gggagctgct	tcttgaggac	ggataacaatg	tttaccagtc	cgaggcccac	420
ggcctcccg	tcacactg	ccccacacc	gggaccctgc	accccgagga		480
ccagctcg	tcctgcact	accaggc	ccccctgcac	ccccagagcc	ccccggaaatc	540
ctggccccc	agcccccg	tgtggctcc	tggaccctc	tgagcatggt	gggaccttcc	600
cagggccgaa	gccccagcta	cgcttcctga				630

<210> SEQ ID NO 163

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Procavia capensis

<400> SEQUENCE: 163

atggactgg	ccaagtttg	gatcgagcac	ccgggactgt	gggtcccgt	gatggcagta	60
cttctgtgg	gagcctg	aggataccct	atccctgact	ccagccccct	tctccaattc	120
ggaggccagg	tccggcaac	g	ttac	ac	cgcaggagac	180
ctggagatcc	gagcagac	gg	cacgggtgt	ggggctgccc	accggagccc	240
ttggagatca	aagcttgc	aa	gcccggata	attcagatct	tggagatcaa	300
ttcctctgc	agggtc	tgtgg	ctg	tatggatcgc	tccgtttga	360
tgcagcttcc	gggagctgct	tcttgc	aa	gat	tggcccac	420
ggcctcccg	tccg	ctacc	atcc	ca	ggac	480
ccagcccg	tcctgc	act	gccagg	ctc	acgg	540
ctggccctg	agcccccg	tgtggactcc	tca	gac	ccatggt	600
cagggccgaa	gccccagtt	cgcttcctga				630

<210> SEQ ID NO 164

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Cavia porcellus

<400> SEQUENCE: 164

atggactgg	cccg	gact	g	gtgtg	gagc	60
-----------	------	------	---	-------	------	----

-continued

tcttcgggtgg gagccctgcaca ggcacacccct atccctgact ccagccccct cctccagttt 120
ggggggccagg tccggcagcg gtacctctac acagatgtg ctcaggacac tgaagtgcac 180
ctggagatca gggccgatgg ctcagttacgg ggcattgccc acaggagccc tgaaagtctc 240
ctggagatca aagccttgc aa gccaggagtc attcagatct tgggaatcag gacttccagg 300
ttcctgtgcc agaggcccg a tggagatctg tatggatcac tccactttga tcctgaggcc 360
tgcagcttcc gggagctgtct gcttgctgat ggctacaatg tctacaagtc tgaagccac 420
ggcctccctc tgcacccgtct gcgcgggtgac tctctatcg c aggaaccagc acccccgagga 480
ccagccccat ttctgcact accaggccctg cccgcacac ccccgagcc acccaggatg 540
ctgccccccag ggccccccaga tgtgggctcc tggacccctt tgagcatggt gggcccttta 600
tgggaccggaa qcccccaccta tacttcctga 630

<210> SEQ ID NO 165
<211> LENGTH: 583
<212> TYPE: DNA
<213> ORGANISM: *Tupaia belangeri*

<400> SEQUENCE: 165

atggggctggg acaaggcccc gttcgagcac ctggggagcgt gggctccctgt gctggctgtc 60
ctcctcttgg gagcctgcga ggcataaccc atccctgact ccagccccct cctacaattc 120
ggggggcagg tccggcagcg gtacacctac acggacgaca cgaggacac agaagccac 180
cttgagatca gggccgacgg cacccgtggtg gggggccccc accaaaagccc ggaaagtctc 240
ctggagatcg aagccttcaa gccgggggtc attcaaatcc tgggagtcgaa gaccccccagg 300
ttcctgtgcc agagggcaga cggggccctg tacgggtcgc ttcacttcga ccccgaggcc 360
tgcagcttcc gggagctgtc tctcgaggat ggataacaaca ttaccaggc tgaggtcg 420
ggccctcccccc tgcgcctgcc gccccacac tccccacatc gggaccggac ccctcgaaaa 480
ccagctcggtt tcctgcccgt gcctggccctg cccctgggtc ctccagagct gccaggggtc 540
ctggggcttgg aaaaaaaaaaa catggactcc tcaagacccggc tga 583

<210> SEQ ID NO 166
<211> LENGTH: 586
<212> TYPE: DNA
<213> ORGANISM: *Sorex araneus*

<400> SEQUENCE: 166

atggctctggg acaaggccag ggggcagcag ttgggactgt gggccccat gctgctggc	60
ttgctgtgg gtgcctgcga ggcacacccc ctccctgact ccagccccct cctccaaattt	120
ggggggccaag tccgactgag gttccctgtac accgacgatg cccagaggac aggggcgcac	180
ctggagatca gggccgacgg cacagtgcag ggtgcggccc acaggacccc agaatgtctc	240
ctggagctga aagccttcaa gccaggcgta attcaaattcc ttggggtcag cacatccaga	300
tccctgtgcc agcggcccgaa tggggtctgt tatggatcgc ttgcgttga cccagaggcc	360
tgcagttcc gggaaacttct tctccaggat ggatataacg tttaccagtc tgaggccctg	420
ggtgtcccgc tctacctaca cccgccccagt gccccagtgt cccaggaacc agcctcacgg	480
ggcgcgcgtcc gtttcctgcc actgccagga ctgcacacccctg cctccctggaa gccccccagg	540
ccccccqccccc cqqtqcctcc aqacqtqqqt tcctcaqacc ccctqa	586

<210> SEQ ID NO 167

-continued

<211> LENGTH: 546
<212> TYPE: DNA
<213> ORGANISM: Ictidomys tridecemlineatus
<400> SEQUENCE: 167

atgtacccca tccctgactc aagccccctc ctccaatttg ggggccaagt ccggcagcgg	60
tacctgtaca cagatgtatgc ccaggagact gaggcccacc tggagatcag ggctgtatggc	120
accgtggtgg gggctgccc tcaaagcccg gaaagtctct tggaaactgaa agccttgaag	180
cctggggtca ttcaaatactt ggggtcaaa acatccagg ttcgtgtccca gagggcagat	240
ggagtgtgt atggatcgct ccactttgac cctgaggccct gcagcttccg ggagcagctt	300
ctggaggacg ggtacaacgt ttaccagtca gaatcccacg gcttccccgt ggcctgccc	360
cctaactcac cataccggga cccagcgcgc ccaggaccag cccgcttct tccactgcca	420
ggcctgcccc cagcagccct ggagccgcca gggatctgg gcccctgagcc ccctgtatgt	480
ggctctcccg acccacttag catggtgggg ccttgcagg gccgaagccc cagttacgct	540
ttctga	546

<210> SEQ ID NO 168
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: Loxodonta africana
<400> SEQUENCE: 168

atggactggg ccaagttgg gttggagcac ccaggactgt gggccctgt gatggctgtc	60
cttctgtgg gagcctgcca gggacaccccc atccctgact ccagccccc cctccattc	120
ggggggccagg tccggcaacg ttacctctac acagatgtac aggagaccga gcccacactg	180
gagatcagag cagatggcac agtggcgaaa gccgctcacc ggagctctga gagtctttt	240
gagctgaaag ctttgaagcc tggataattt cagatctttt gggtaagac atcccggttc	300
ctgtgccagg ggcctgtatgg ggtgtgtac ggatcgctcc atttcgaccc agccgcctgc	360
agttccggg agctgtttt tgaagatggg tacaatgttt actggccga gcccatttgg	420
ctcccaatcc gctgtccctc ccacaactcc ccatataggc acccagcatc ccgggtacca	480
gcccgtttcc tgccactgcc aggcctgtcc ccaatgtcc aagaacctcc aggggtcttg	540
gcccctgagc cccctgtatgt ggactcctca gacccctga gcatggtggg gccttcacag	600
ggccgaagcc ccagctatgc ttctgt	627

<210> SEQ ID NO 169
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: Sus scrofa
<400> SEQUENCE: 169

atggggctggg ccgaggccaa gttcgagcgc ttgggactgt gggccctgt gctggctgtc	60
ctgctgggag cctgcccatt cctgactcca gccccctctt ccaattttgg	120
ggcccaagtgc gccaacgata cctctacacg gatgtgccc aggaaactga agcccacctg	180
gagatcagag ctgtatggcac cgtggcagggt gtagccgc agagccctga aagtctttt	240
gagctgaaag ccctgaagcc aggggtcatt caaatggg gatgtccagac atcccggttc	300
ctgtgccagg ggccagacgg gagactgtac ggatcgctcc atttcgaccc tgaggccctgc	360
agttccggg agctgtttt tgaggatggc tacaacgtttt accagtctga ggcccttggc	420
ctcccaatcc gctgtccctc gcaccgcgtcc tccaaacgggg acctggccccc ccggggaccc	480

-continued

gctcgcttcc tgccactgcc aggcctgccc ccggcacccc cggagccgcc agggatcttg	540
gccccctgaac ctcccgacgt gggctcctcg gacccctga gcatggtggg gccttcacac	600
ggccggagcc ccagctacac ttcttga	627

<210> SEQ ID NO 170
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: Felis catus

<400> SEQUENCE: 170

atggggctggg acgaggccgg gtcccagcgc ctgggactgt gggtcgtgct ggggttcctt	60
ttgccggaag cctgecaggc acaccctatac cctgactcca gccccctcct ccaattcggg	120
ggccaagtcc gacagcggtt cctctacacg gacgacgccc aggagacaga ggtccacctc	180
gagatcaagg ctgatggcac agtgggtgggg accgctcgcc ggagccctga gagtctttg	240
gagctaaaag ccctgaagcc ggggttaatt caaatcttgg gggtaaaaac gtccagggttc	300
ctgtgccagg gcccagatgg gacactgtat ggatcgctcc gtttgaccc cgccagctgc	360
agcttcgggg aactgtctt ggaggacgga tacaacatct accactcgga gaccctcggg	420
cctccactcc gctgtcccccc ccacaactcc ccataccggg acttggccc cggggcacct	480
gcccgettcc tgccgctgcc aggcctgttt cccgcacccc cggagectcc agggatctg	540
gccccggagc ccccgacgt gggctcctcg gaccctctga gcatggtggg gccttccag	600
ggccgaagtc ccagctacgc ttcttga	627

<210> SEQ ID NO 171
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: Otolemur garnettii

<400> SEQUENCE: 171

gacaaggcca ggactgggtt caagcacca ggaccatggt ttcccctgct ggctgtactt	60
ttgttgggg cctgecaggc acaccctatac cctgactcca gccccctact ccagtttgtt	120
ggccaagtcc ggcagcggta cctctacaca gatgatgccc aggagacaga agccacactg	180
gagatcaggg aagatggcac agtgggtgggg gctgcacaac agagccctga aagtctttt	240
gagctgaaag cttaaagcc aggggtcatt caaatcttgg gagtcaagac atccagggttc	300
ctgtgccaga gcccagatgg gggcttatat ggatcgctct actttgaccc caaggcctgc	360
agtttccggg agctgtttt tgaggatgga tacaacgttt actggcttga gacctatggc	420
cctccactgc acctgcctcc tgccaaattcc ccatactggg gcccattccct tcggagccca	480
gcccgettcc tgccgactgcc aggcctctt gcagcatccc cagagctgcc ggggatctg	540
gcccgttggaa ccccccgtgtt gggctcctcg gaccctctga gcatggtggg gccttcgcag	600
ggccgaagcc ccagctatgc ttcttga	627

<210> SEQ ID NO 172
<211> LENGTH: 627
<212> TYPE: DNA
<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 172

atggactgga tgaaatctag agttggggcc ccgggactgt gggctctgtct cctgctgcct	60
gtcttcctgc tgggggtgtg cgaggcatac cccatctctg actccagccc cctccatccag	120

-continued

tttgggggtc aagtccgaca gaggttatctc tacacagatg acgaccaggaa caccgaagcc	180
cacctggaga tcagggagga cggaacagtgc gtgggcacag cacacccggc tccagaaaat	240
ctcctggagc tcaaaggctt gaagccagggt gtcattcaaa tcctgggtgt caaaggctct	300
aggtttttt gccaacaacc agatggaaact ctctatggat cgcctcactt tgatcctgag	360
gcctgcgtt tcagagagct gctgcttaag gacggataca atgtgtacca gtctgaggcc	420
catggeetgc ccctgcgtct gccccagaag gactcccagg atccagcaac ccggggaccc	480
gtgcgttcc tgcccatgcc aggcctgccc cacgagcccc aagagcaacc aggagtctt	540
ccccccagagc ccccgatgt gggttcctcc gacccctgaa gcatggtaga gccttgcaa	600
ggccgaagcc ccagctatgc atcttga	627

<210> SEQ ID NO 173

<211> LENGTH: 633

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 173

atggaatggta tgagatctag agttgggacc ctgggactgt gggtccgact gctgctggct	60
gtcttcttcgc tgggggtcta ccaagcatac cccatccctg actccagcccc cctccctccag	120
tttgggggtc aagtccggca gaggtacctc tacacagatg acgaccaaga cactgaagcc	180
cacctggaga tcagggagga tggAACAGTG gtaggcgcag cacccggcag tccagaaaat	240
ctcctggagc tcaaaggctt gaagccagggt gtcattcaaa tcctgggtgt caaaggctct	300
aggtttttt gccaacagcc agatggagct ctctatggat cgcctcactt tgatcctgag	360
gcctgcgtt tcagagaact gctgctggag gacggttaca atgtgtacca gtctgaagcc	420
catggeetgc ccctgcgtct gcctcagaag gactccccaacc accaggatgc aacatccctgg	480
ggacccgtgc gtttctgccc catgccaggc ctgctccacg agccccaaga ccaagcagga	540
ttccctgcccc cagagcccccc agatgtgggc tcctctgacc ccctgagcat ggttagagcc	600
ttacagggcc gaagccccag ctatgcgtcc tga	633

<210> SEQ ID NO 174

<211> LENGTH: 324

<212> TYPE: DNA

<213> ORGANISM: Vicugna pacos

<400> SEQUENCE: 174

atggactggg acgaggccaa gttcgagcat cggggactgt gggtcccagt gctcaactgtc	60
cttctgtctgg gaggcctgcca ggcacgcccc attcctgact ccagccccct cctccaaattc	120
ggggggccaaat tccggcagcg gtacctctac acggatgcacg cccaggagac agaagccccac	180
ctggagatca gggctgtatgg cacagtggtg ggggtggccc gccagccccga aggaattcct	240
cccgagccctc ctgacgtggg ctcctcagac ccctgagca tgggtggggcc ttccatacagc	300
agaagccccca gctacacttc ctga	324

<210> SEQ ID NO 175

<211> LENGTH: 633

<212> TYPE: DNA

<213> ORGANISM: Anolis carolinensis

<400> SEQUENCE: 175

tgtaaaagca agggaggagg gaagggggga gagaggatgt gggttagaccc agttttctgg	60
gtgccttgc tccgcacagc tcctgctttt ccctgcggaa attccaaaccc catctaccaa	120

-continued

tttgatggc aggtccggct tcggcaccc tacacagcag atgaacagac gcacccac	180
ttggagatct tgccagacgg taccgtgggt ggatccagggt ttcaaatcc cttcagtttgc	240
atggagatca aagctgtgaa gccaggaggc attcgcatgc aggccaagaa gaccttaga	300
tttctctgtatgaaacccaa tggacgactg tatggctcgc tgttctactc tgaggaggca	360
tgcaacttcc atgagaagggt tctcagcgat ggctacaacc tctactattc tgaaaactac	420
aacatacctg tcagectcag ctccggcagg aacctgggtc agagccgtca gttgcctccc	480
ttctccaaat tcctgccgtt agtcaacaaa attccttttgc agcctgtgt tgaagacttt	540
gacttctatg gacatcaatt ggtatgttgc tcaagctgtatc cttttagcat ttttaggacaa	600
aaccctgggtt tcatgagtcc gagctatgtc ttc	633

<210> SEQ ID NO 176

<211> LENGTH: 564

<212> TYPE: DNA

<213> ORGANISM: Gadus morhua

<400> SEQUENCE: 176

ctccctctcg ccacccctccct ccacatccgc ctctcccttc acgtccccga ctccggccccc	60
ctgctgtggc tggggcggacca ggtcaggaggag agacacccctt acacagcaga gagccacccgg	120
agggggctgt tcctggagat gageccggac ggtcagggtga caggaagtgc tgctcagacg	180
cggctcagtg ttctggagct gaggtccggc agagcaggag atacggtcat cagagcgcgc	240
ctctcccttc tctacctgtg tggacagg gcaggtcacc tgacaggaca gagacagtac	300
acagagtccg actgeacccctt cagagaggc atccttggagg acggctacac ccacttccgt	360
tccgtgcacc acggacttcc tatttcgtgc ggcggcggac actccccagg gagacagggg	420
ctgcgttca gcagggttccct cccgctgagg agcagtctgt cagaggatag ggtcgccgag	480
ccccccagaca gcccactgaa cctggactct gaagaccccc tggggatggg tctgggttcg	540
ctccctcagcc cggcccttc catg	564

<210> SEQ ID NO 177

<211> LENGTH: 648

<212> TYPE: DNA

<213> ORGANISM: Latimeria chalumnae

<400> SEQUENCE: 177

atgttatgcc agagtttgtt gatattaagt cagaaattca tttttggct cttttgtact	60
ggattggggc taacaggatt ggcttgaca aggcccttcc aggattccaa tcccatctg	120
cagtattccg attccatccg gctccgacat ctgtacactg ccagtggagag tcggcacctt	180
cacctacaaa tcaactccga tggacagggtgg ggagggacaa ccaagcaag cccttacagt	240
ctgttggaga tgaaggcggtt gaagacagggt tttgtggta tcaggggacaa gaaaagcgcc	300
cgttacctt gtatggAACG tagtggacgg ctctatggat cgctgcagta tacagaaaa	360
gactgcaccc tcaaaggagggt tggatggca gatggataca acctgtatgt ctcagaggaa	420
caccaggcaca cagtgcacgtt gagcccccattt agggcgaggaa tagcgcacagg gaaaaagatc	480
ccacccctttt cccatcccttccatggatg aacaagggtgc ctgtggaggta tggccgt	540
gagatggaggat ttgtccagggt gctgcgggaa atgacggccgc acgtggactc tccggatccc	600
tttggatgaa cctggaaaga atcggttccat agtccggactt tttttggcc	648

<210> SEQ ID NO 178

-continued

<211> LENGTH: 579
<212> TYPE: DNA
<213> ORGANISM: *Tursiops truncatus*
<400> SEQUENCE: 178

atgggctggg acaagaccaa actcgagcac ctggactgt gggccctgt gctagctgc	60
ctgctggac cctgccagc acatcccatt cctgactcca gccccctctt ccaatttggg	120
ggccaagtcc gccagcgata cctctacacg gatgacgccc aggagacgga ggcccacctg	180
gagatcaggg ctgatggcac agtggtgggg acggcccgcc ggagcccgaa aggagttaaa	240
acatccaggat tcctgtgcca gggccagag gggaggctgt atggatcgct ccacttcaac	300
ccccaggcct gcagcttcg ggagctgctt cttgaggatg gataacaacgt ttaccagct	360
gaggctttg gcatccctt ccgcctgccc cgcacccgt cctccaactg ggacctggcc	420
ccccggggac ctgctcgctt cctgcccgtg ccaggcttcc tcccgcacc cctggaggct	480
ccagggatct tggcccccga gcctccaaac gttagttccct cggacccctt gagcatggtg	540
ggacccctcac atggccgaag cccagctac acttcctga	579

<210> SEQ ID NO 179
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: *Mustela putorius furo*
<400> SEQUENCE: 179

atgggctggg aagaggccag gtccgagcac ctggggctgt gggccctgt gctggcggtc	60
cttttgcgg gggctgcca ggcataccctt attcctgact ccagcccccctt cctccatattt	120
ggaggccaaat ttcgacagcg gtaccccttac acagacgacg ctcaggagac ggaggccac	180
cttagatca gggctgtatgg cacgggtggg gggctgccc gcccggccccc cgaaaatctc	240
ttggagctga aagccctgaa gcccgggtc attcagatct tgggagtgaa aacatccagg	300
tccctgtgcc agggcccgaa tgggacactg tacggatcgt tccacttgc cccctgttagcc	360
tgcagcttcc gggaaatgtct tctggaaatggat ggataacaaca tctaccactc tgagaccctg	420
ggccctccac tgcgcctgccc ccccccacaac tccccacaca gggacctggc gccccgggg	480
cctggccgtt tcctggccctt gcccggctg cttccggcca ccccgagtc cggggggatc	540
ccagcccccg agcctccaaat cgtgggctcc tcaagacccccc tgagcatggt ggggccttg	600
cagggtcaaa gtccctagcta cacttcctga	630

<210> SEQ ID NO 180
<211> LENGTH: 597
<212> TYPE: DNA
<213> ORGANISM: *Takifugu rubripes*
<400> SEQUENCE: 180

tttatttttatttttttca aactgcactt tttccctt ccaaattggtt caacttttat	60
ctccctgact ccaacccgctt cttatccctt gacagtcatg gcagaggcat ccacccctac	120
acagataatc aaaggcgagg gatgtatctg cagatgacca cagatggaaat cgttccggg	180
agtgtatgtcc agacggcgaa cagtgtgtcg gaaatgtatgt cagtcgaaaa cggccacgtc	240
gtcatccgag gaaaatcgatc ttctctgtttt ctctgtatgg acagcagagg ccgtttatgg	300
ggccagggc accccactga ggcggactgc actttcaggag aagtgtgtcc ggcagatggaa	360
tacactcgct tcctgtccctt gcacaacggaa actccctgtgt ctctggacc taaacaatct	420
ccagaccaggc acacagttcc cttctactgtt tcctgtccgc tcaggaatac actggcagag	480

-continued

gagagcatgt ctgaaccacc atcaaaccaa cagagatatt ttaacattga ctctgatgat	540
cttcttgaa tggattaaa tgcatggtc agtcctcagt ttcaggaga caagtga	597

<210> SEQ ID NO 181
<211> LENGTH: 612
<212> TYPE: DNA
<213> ORGANISM: Dipodomys ordii

<400> SEQUENCE: 181

atggaccagg caaagaccag gttggggcc cggggactgg ggggcattgt gctggctgca	60
ataattctgg gagcatgcaa ggcacggct atccctgact ccagccccct cctccaattt	120
gggggtcaag ttccggcttcg gcacctctac acagatgaca ctcaggagac ggaagccat	180
ctggagatca gggcagatgg cacggtagtg gggactgccc accggagccc tgaaagtctc	240
ttggagatca aagccttgaa gccaggagtc attcaaatct tagggatcaa gacatccaga	300
ttcttatgcc agagaccaga cgggacactg tatggatcac tccacttga ccctgagggt	360
tgcagttcc aggagatgct tcttggaaat ggataacaaca tttaccgttc tgaaggccctg	420
ggtctccccc tgcgectgca cccagatcca gcaccctggg ggccagcccg ctccctgccc	480
ctgcctggtg tgccccccgc accggccggag cccccccggg taatggatcc cgaacccct	540
gatgtggct ctcggaccc tcttggatgt gtgggactgt tgcaggcccg aagccccagc	600
tatgcacccct ga	612

<210> SEQ ID NO 182
<211> LENGTH: 555
<212> TYPE: DNA
<213> ORGANISM: Echinops telfairi

<400> SEQUENCE: 182

atgggttgca ccaaattctgg gtggaaatggcc cggggactgt gggccctgt gctggccagc	60
tttctgtgg gaggctgccc agcacacccc atccctgact ccagccccct cctccaattt	120
gggggcaag tccggcagcg ataccttat acggatgacg cccagaccac cgaggccac	180
ctggagatca gagcggatgg cacagtgggg ggcgtcgccc accagagccc agagaaggcc	240
ctgagtaat ggctgtaaaa gcccctgaga taactccatt tgcaccacgc cgcctgcagc	300
ttccgggaga agcttctaga agacggatac aacttgtacc actctgagac ccacggctc	360
ccctccggcc tcccaccccg tggggccgac ccctttctc agcctggggc cgcctccca	420
ccgctggccg gccagctccc acaactccaa gagacgcccag gggccctcgc ccccgaaacc	480
cccgacgtgg gctttcaga cccctgagc atggggggc cttggcgagg gcaaagtccc	540
agttatgcct cctga	555

<210> SEQ ID NO 183
<211> LENGTH: 633
<212> TYPE: DNA
<213> ORGANISM: Macaca mulatta

<400> SEQUENCE: 183

atggactcgacgagacccgg gttcgagcac tcaggactgt gggccctgt gctggctgg	60
cttctgtgg gagcctgcca ggcacacccc atccctgact ccagtcctct cctgcattt	120
ggggggcaag tccggcaacg gtacctctac acagatgtat cccagccagc agaagccac	180
ctggagatca gggaggatgg gacagtgggg ggcgtgttc accagagccc cgaaagtgg	240

-continued

tgtgggccag	agcctgggtc	tgagggagga	ggggctgtgg	gaggtgctga	gggacctgga	300
ctcctgggtc	tgagggaggc	agggctgggg	cctggatcc	ggctccactt	tgaccctgag	360
gcctgcagct	tccggagct	gcttctttag	aacggataca	atgtttacca	gtccgaggcc	420
cacggectcc	cactgcacct	gccgggaaac	aagtccccac	accgggaccc	tgcattccaa	480
ggaccagctc	gcttcttgc	actaccaggc	ctggccccc	caccccccgg	gccgcagga	540
atcctegccc	cccagccccc	cgatgtgggc	tcctcgacc	ctctgagcat	ggtgggacct	600
tcccaggccc	gaagccccag	ctatgcttcc	tga			633

<210> SEQ ID NO 184

<211> LENGTH: 582

<212> TYPE: DNA

<213> ORGANISM: *Microcebus murinus*

<400> SEQUENCE: 184

atgggctggg	acgaggccgg	cggccgggttc	gagcacccag	gactgtggtt	tcccatgctg	60
ggtgtctgc	tgctgggagc	ctgcccaggcg	taccccatcc	ctgactccag	ccccctctc	120
caattttgcg	gcacaagtccg	gcagcggcac	ctctacacag	acgatatcca	ggagacagaa	180
gcccacctgg	agatcagggc	ggacggcaca	gtggtggggg	ccggccgaca	gagccctgag	240
ttggagctga	aagccttaaa	gccaggggtc	attcaaatct	tgggagtcaa	gacctccagg	300
ttcctgtgcc	agaggccaga	cggggccctg	tacggatcgc	tccacttga	ccccgagtgc	360
agttccggg	agctgtttct	tgaggatgga	tacaacgtct	actgtcccta	cctcccgctg	420
cacctgtccc	cacgcacatcg	actggccgga	tcacgctctg	cgctgccact	gcccccagca	480
cctgaacgca	ggattttggc	cccgagggcc	ccggatggct	cctcgacccc	tctgagcatg	540
gtggggcctt	cgcaggcccg	aagtcccaga	tatgcttct	ga		582

<210> SEQ ID NO 185

<211> LENGTH: 615

<212> TYPE: DNA

<213> ORGANISM: *Ochotona princeps*

<400> SEQUENCE: 185

aaagacatgg	acgggctcca	gcctccgggg	ctgcgggttc	ctgtgtggc	tgcctgtct	60
ttgggagttg	gccaggcacg	ccccatccct	gattctagcc	ctctccatca	attcgggggc	120
caggtccggc	agaggcacct	ctacacggat	gacgcccagg	aatcggaaat	acacactggag	180
atccgggacg	acggcacctgt	ggcaggggact	gcccggccga	gccctgaaag	tctcttagaa	240
atgaaagcgt	tgaagccagg	cgtcatttag	atcttggggg	tccacacatc	caggttctg	300
tgccagagac	cacacgggac	gctgtacggc	tcgttccact	tgcaccacaa	ggcctgcagc	360
ttccgggagc	agctgttgaa	ggatgggtac	aacgtgtacc	actcagagac	acacggcctc	420
cgcgtgcgcc	tgtctccaga	ccgagccccc	cgggggccag	cccgcttct	gccactgcca	480
ggcccttcctc	ctgacccct	ggtgccaccc	ctgccaccgg	acgttccatc	ccctgagccc	540
cccgacgtgg	actccccaga	ccccctgac	atggtggggc	ccttgcaggg	ccaaagcccc	600
agctacactt	cctga					615

<210> SEQ ID NO 186

<211> LENGTH: 546

<212> TYPE: DNA

<213> ORGANISM: *Xiphophorus maculatus*

<400> SEQUENCE: 186

-continued

tgcgggttcc	ccttcctttt	cttaatcctc	tctcttcctt	ttttctcttc	ctcgttttac	60
atccccagaat	ccaacccaat	ctttgcccttc	aggaatcagc	tcagagaggt	gcatctctac	120
acagaaaatc	acagacgggg	tttgttatgtg	gagatacata	tggatgggag	agtgaactgga	180
agtgtatgctc	agagtcccta	tagtgtgttg	cagataaaagt	ctgttaaacc	gggtcatgtg	240
gtcataaaagg	gacagacatc	gtccctgttc	ctctgcatgg	acgactccgg	aatctaaga	300
ggacagacaa	cctatgacga	ggctgactgc	tccttcagggg	aactgtgtct	ggccgatggc	360
tacacccgtt	tcctgaactc	acaacatggc	gttcctttat	cactggcata	cagaaaactct	420
ccagatcgac	actccgttcc	tttcacaaga	tttttacctc	tcaggaatac	tttaacggtt	480
tcagaagaat	caacaaaaac	tcagagggac	ttcaacctgg	actcggacga	ccttctcggt	540
atggga						546

<210> SEQ ID NO 187

<211> LENGTH: 576

<212> TYPE: DNA

<213> ORGANISM: Gasterosteus aculeatus

<400> SEQUENCE: 187

tctctctcc	tcatggtccc	acttccttcc	tgttcatcc	tttatctcac	tgactccagc	60
ccacttctac	ccttcaataa	tcaagtcaaa	gaggtgcacc	tctacacagc	agagaatcac	120
agaagagcga	tgtacctgca	gategctctg	gacgggagcg	tgtcgggaag	cgacgctcg	180
tccacttaca	gtgtgctgca	gctgaaatct	atccagccgg	gccacgtgg	catcagaggg	240
aaggcctcct	ccatgttcc	ctgegtggac	agcgggggccc	gtttgagagg	acagggggcg	300
tactcagagg	ccgactgcag	cttcaggggag	ctgctgtctgg	gggatggcta	caccgggttc	360
ctgtcctcgc	agcaegggtc	cccgctgtct	ctggcgatcg	ggccttcccc	ggatccaaac	420
tcgggtccct	tcaactcgatt	cctaccatc	cgaccgc	ccgaggctga	gagcgtgatc	480
gaagagccac	cgagcaatca	gagatacgtc	aacgtggact	ccgaggatct	tcttggaaatg	540
ggcctgaaca	ctgtggtag	tcctcagttc	tcggcg			576

<210> SEQ ID NO 188

<211> LENGTH: 495

<212> TYPE: DNA

<213> ORGANISM: Sarcophilus harrisii

<400> SEQUENCE: 188

gtgtctgcca	tgggcctgag	ggagcgagct	cccaggtacc	tggcccccgt	gctgtccctg	60
ctcttggcct	gcagggcctc	gggtcacc	ctcccgatt	ccagccccat	gctctgttt	120
ggggggcagg	tccgcctccg	gcacccctac	acggatgtgg	gccaggaggc	cgaggccac	180
gtggaaactgg	cgtccgacgg	cacagtccgg	gccccagcgc	ggaggagatcc	caacagtctc	240
ctggagctga	aggctgtgaa	gccgggcata	gtccgaatcc	tggccgtcca	cagctctcg	300
tttctgtgta	tgaggccaa	cggggagctg	tacggagcga	tacactacga	cccttccgc	360
tgcaacttc	gggagcgct	gctggggac	ggotacaac	tgtacgagtc	cgaggctcac	420
gggaggaccc	tccgcctgac	ccccaaaggcc	gcacccggac	ccgcccggacc	ttctcgcttc	480
ctgcccgtcc	ccggc					495

<210> SEQ ID NO 189

<211> LENGTH: 627

<212> TYPE: DNA

-continued

<213> ORGANISM: Macropus eugenii

<400> SEQUENCE: 189

acagaggagc	cttctactgg	gtccaggcac	ctgggacaat	gggcctccgg	gctgectgg	60
cctctgtgt	cttgtctct	ggcctacagg	ggctggggct	ccccatccc	tgattccagc	120
cccatgtcc	tgttgtgtgg	ccaggtccgc	ctccgacacc	tgtacacaga	tgtggcag	180
gacacggagg	cccatgtgga	gctggggeca	gatggagtgg	ttcgagctgt	ggctgagagg	240
agccccaaaca	gtcttctgga	actgaaggcg	gtgaagcctg	gagtcatccg	aatcctcgct	300
gtccagagct	ctcggttct	gtgttatgagg	cccaacgggg	aactgtatgg	agcggtaaac	360
tatgaccctt	ctgcctgcaa	cttcgggaa	catctgtgg	gggatggta	taatgtgtat	420
gaatcagaga	ctcacagaag	gaccctccgt	ctgtccccc	ccctgggtca	ggctggcccc	480
tctcggttcc	tgccacttcc	aggcgactgg	ctgccccggcc	ctgateccacc	ttggggcacag	540
ggccctgagc	ccccagacgt	gggctctgca	gaccctctga	gcatggtggg	ggccgtgcag	600
ggcctcagcc	ccagctactc	ctcctga				627

<210> SEQ ID NO 190

<211> LENGTH: 480

<212> TYPE: DNA

<213> ORGANISM: Xenopus tropicalis

<400> SEQUENCE: 190

agagggggta	ggacaaaaaa	aaagacgtta	ctcaggaaat	ggctttgcct	tttagccatt	60
atgtttagta	ggtaagggtt	ttcttttagca	aatcctatcc	agaattcgaa	cccaatctta	120
tccaaacgaca	accaagtacg	gactcagtt	ttatacacag	ataacaataa	catgcacctg	180
tatcttcaga	tcacccacaa	tggagtagta	actggtacgg	aagaaaagaa	tgactatgg	240
gtgctggaaa	taaaggcagt	aaaagctggg	gtttagtta	taaaaggaat	tcgaagcaat	300
ctctacatat	gcatggattc	tagacacaa	ttgtatgcgt	cgccatata	taaagatgac	360
tgccatttcc	atgaaaagat	cacaccagat	aattacaaca	tgtatagctc	agagaagcat	420
tcagaatacg	tgtccttagc	tccattaaaa	ggaagccaga	tggctcg	tttctacata	480

<210> SEQ ID NO 191

<211> LENGTH: 585

<212> TYPE: DNA

<213> ORGANISM: Danio rerio

<400> SEQUENCE: 191

atgcttcttg	ctctgttttt	tatatttttt	gtcttttttc	ctcatcttcg	gtgggtatg	60
tatgttcctg	cacagaacgt	gttctgcag	tttggcacac	aagtcaaggaa	acgcctgctt	120
tacacagatg	ggttgtttct	tgaaatgaat	ccagatggct	ccgtcaaagg	ctctctgtaa	180
aagaatctaa	atttgtgtct	ggagctgcgt	tcaatcaag	cggggtgaaac	cgtcatccag	240
agtgcagcta	catctctcta	cctctgcgtc	gatgtcaag	acaagctgaa	aggacagcat	300
cattactctg	cactagactg	caccttcag	gaattgtac	tggatggata	ttcggttttc	360
ctttctccac	acactaatct	tccctatcg	ctcctctcg	aacgtcagaa	acacggcaat	420
cctctttctc	gttctctccc	tgttagcaga	gcagaggaca	gccggacaca	ggaggtgaaa	480
cagtatattc	aggatataaa	cctggactct	gacgaccac	taggaatggg	acatcggtca	540
cacttacaga	ccgtcttcag	tcccagtcg	catactaaaa	aatga		585

-continued

<210> SEQ ID NO 192
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Bos grunniens mutus

<400> SEQUENCE: 192

atgggctgg	atgaagcgaa	atttaaacat	ctgggcctgt	gggtgccggt	gctggcggtg	60
ctgctgtgg	gcacctgccc	cgcgcaccc	atcccgata	gcagccccgt	gctgcagttt	120
ggcggccagg	tgcgecagcg	ctatctgtat	accgatgtat	cgcaggaaac	cgaagcgcat	180
ctggaaatcc	gcgeggatgg	caccgtggtg	ggcgccggcgc	gccagagccc	ggaaaggctg	240
ctggaaactga	aagcgctgaa	accgggcgtg	attcagattc	tggcgctgaa	aaccagccgc	300
tttctgtgcc	agggccccgga	tggcaaactg	tatggcagcc	tgcattttga	tccgaaagcg	360
tgcaagtttc	gcgaactgtct	gctggaaat	ggctataacg	tgtatcagag	cgaaacctg	420
ggcctgcgc	tgcgectgcc	gccgcagcgc	agcagcaacc	gcatccggc	gccgegcggc	480
ccggcgccgt	ttctgcgcgt	gccgggcctg	ccggcggaac	cgcggatcc	gccgggcatt	540
ctggcgccgg	aaccgcggaa	tgtggcagc	agcgatccgc	tgagcatggt	ggggccggagc	600
atggccgca	gccccagcta	taccagctaa				630

<210> SEQ ID NO 193
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Saimiri boliviensis boliviensis

<400> SEQUENCE: 193

atgggctcg	aggaggtcgc	gttggagegc	cctgcactgt	gggtctctgt	gttggctgg	60
ctcctgtgg	gaacctgcca	ggcataaccc	atccctgact	ctagtcctct	cctgcattt	120
ggaggccaag	tccggcagcg	gtacctctac	acagatgacg	ctcagcagac	agaagccac	180
ctggagatca	gggaagatgg	cacggtgccg	ggggctgcc	accagagccc	cgaaagtctc	240
ttgcagctga	aagcattaaa	gccagggtt	attcaaatct	tggagatcaa	gacccagg	300
ttcctgtgcc	agaggccgga	cggggccctg	tacggatcgc	tctactttga	ccccgaggcc	360
tgcagettcc	gggagctgtct	tcttgaggac	ggataacaatg	tgtaccagtc	cgtggccac	420
agcctccgc	tgcacactgcc	agggggcagg	tccccaccc	gggacectgc	acctcgagga	480
ccagctcgct	tcctgcgcgt	accaggccctg	ccccccgaac	cccccgaggc	gccaggaatc	540
ctggcccccg	agccccccga	tgtggcctcc	tcagaccctc	tgagcatggt	ggggccttcc	600
caaggccaaa	gccccagcta	cacttcctga				630

<210> SEQ ID NO 194
<211> LENGTH: 633
<212> TYPE: DNA
<213> ORGANISM: Callithrix jacchus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (190)..(190)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 194

atgggctcg	aggaggtcgg	gttggagcac	cctgcactgt	gggtttctgt	gctggctgg	60
ctcctgtgg	gaacctgcca	ggcgcacccc	atccctgact	ccagtcctct	cctgcattt	120
ggaggccaag	tccggcagcg	gtacctctac	acagatgacg	cccagcagaa	agaagccac	180
ctggagatcn	aggaagatgg	cacagtggcc	ggggctgcca	ccaaagtccc	gaaagtgagt	240

-continued

ctcttcagc tgaaaggcctt aaagccaggg gttattcaaa tcttgggagt caagacatcc	300
aggttcttgt gccagaggcc agacggggcg ctgttatggat cgctccactt tgaccccgag	360
gcctgcagct tccggagct gcttctttag gacggataca atgtgtacca gtctgtggcc	420
cacggcctcc cgctgcacct gccagagac aggtcaccac cccgggaccc tgcaccccgaa	480
ggaccagctc gcttcttgcc actaccaggg ctgccccctg aaccccccaga gccgcagga	540
atcctggccc ctgagcccc cgaegtgggc tectcagacc ctctgagcat ggtggggcct	600
tcccaaggcc aaageccccag ctacgcttcc tga	633

<210> SEQ ID NO 195

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Tupaia chinensis

<400> SEQUENCE: 195

atgggctggg ataaaagcgcg ctttgaacat ctggggcggt gggcgccgggt gctggcggtg	60
ctgctgctgg gcgcgtgcca ggcgtatccg attccggata gcagccccgt gctgcagttt	120
ggcggccagg tgcgcgcaggc ctatctgtat accgatgata cccaggatac cgaagcgcat	180
ctggaaatcc ggcggatgg caccgtgggt ggcgcggcgc atcagagccc ggaaagcctg	240
ctggaaactga aagcgctgaa accggggcggt attcagattc tggggcgtgaa aaccagccgc	300
tttctgtgcc acgcgcggg tggcgcgctg tatggcagcc tgcattttga tccggaaagcg	360
tgcagcttcc gcgaactgct gctggaaagat ggctataaca ttatcagag cgaagcgcc	420
ggcctgcgcgc tgcgcctgccc gcccgcattat agcccgcatc ggcgcgcac cccgcaggc	480
ccggcgcgcgt ttctgcgcgt gccgggcctg ccgcgtgggt cgcggaaact gccgggcgt	540
ctggcgctgg aaccgcggg tggggcagc agcgatccgc tgagcatgat gggcccgagc	600
caggggccaga gcccgcgcta tgcgagctaa	630

<210> SEQ ID NO 196

<211> LENGTH: 567

<212> TYPE: DNA

<213> ORGANISM: Papio anubis

<400> SEQUENCE: 196

atggactcgg acgagacccgg gttcgagcac tcaggactgt gggttcctgt gctggctgg	60
cttctgctgg gagcctgcca ggcacacccc atccctgact ccagtcctct cctgcattc	120
ggggggccaaag tccggcaacgg gtacccctac acagatgtat cccagcagac agaagccac	180
ctggagatca gggaggatgg gacagtgggg ggcgcgtctc accagagccc cgaaagtaag	240
tgtggcccaag agcctgggtc tgaggaggaa ggggctctcc actttgaccc tgaggcctgc	300
agcttcggcg agctgcttct tgagaacggg tacaatgttt accagtccga ggcgcacggc	360
ctcccaactgc acctgcgggg aaacaagtcc ccacacccggg accctgcattc ccgaggacca	420
gctcgcttcc tgccactacc aggcctgccc cccgcacccc cagagccacc aggaatctc	480
gcccccccaagc ccccccgtgt gggctctcg gaccctctga gcatgggtggg accttccca	540
gcccgaagcc ctagctacgc ttccctga	567

<210> SEQ ID NO 197

<211> LENGTH: 627

<212> TYPE: DNA

<213> ORGANISM: Pteropus alecto

<400> SEQUENCE: 197

-continued

```

atgggctggg gcaaagcgcg cctgcagcat cccggcctgt gggggccgggt gctggcggtg      60
ctgctggcg cgtgccaggc gcatccgatt ctggataga ccccgctgtt tcagttggc      120
agccagggtgc gcccggccta tctgtatacc gatgatgcgc aggataccga agcgcatctg      180
gaaattcgcg cggatggcac cgtggggcgc gcccggcgcgc gcagcccgga aagcctgtg      240
gaactgaaag cgctgaaacc gggcggtgatt caggtgctgg gcgtgaaaac cagccgcctt      300
ctgtgccaggc gccccggatgg caccctgtat ggcagccgtc attttgatcc ggccggcggtc      360
agctttcgcg aactgctgtct gaaagatggc tataacgtgt atcagagcga agcgctggcg      420
cgccccgtgc gcctggcgcc gtatagcgc cccggcgccg atccggcgccg ccggccggccg      480
gcccgggttc tgccgctgcc gggcccgccg cccggaaaccgc cccggccgcgc gggccgcctg      540
gcccggaaac cggccggatgt gggcagcgc gatccgctga gcatgggtgtg gcccggccgc      600
ggccggcagcc cggatatac cagctaa                                         627

```

<210> SEQ ID NO 198

<211> LENGTH: 633

<212> TYPE: DNA

<213> ORGANISM: Heterocephalus glaber

<400> SEQUENCE: 198

```

atggattggg cgcgcggca aagegaacgc cccggcctgt ggggtgcggc ggtgtggcg      60
gtgtgtgtgc tggggcggtg ccaggcgcat cccatccgg atagcagccc gctgtgtcag      120
tttggggcc aggtgcgcga gcgcacatctg tataccgtat atgcgcaggta taccgaagt      180
catctggaaa ttgcgcggca tggcagcgtg ggcggcgccgg cccatcgccg cccggaaacg      240
ctgtgtggaaac tgaaacggcgt gaaaccggcgt gttgttgcgtt ttctggcggt ggcgcaccagg      300
cgctttctgt gccagcgccc ggtatggcacc ctgtatggca gctgtcattt tgatccggaa      360
ggtgtcagct ttgcgcgact gctgtgtgg gatggctata acatttatca gagcgaagcg      420
tatggcctgc cgctgcgcgt gctgtccggc gatagcgcga gccgcgtatcc ggtgcggccg      480
ggcccgccgc gcttttgcc gctgcggggc ctgcacccgc cccgcgtggaa accgcggggc      540
atgtgtccgc cggaaaccggc ggtgtggggc agcagcgtatc cgctgagcat ggtggggccg      600
ctgcaggggcc gcagccccgag ctatgcgtt taa                                         633

```

<210> SEQ ID NO 199

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Cricetulus griseus

<400> SEQUENCE: 199

```

atggactgaa tggaaatctgg agttggggtc cccggactgt ggggtccctct gctgcctatc      60
ttcctgtgg ggggtctccca ggcacaccccc atccctgact ccagccccct cctccagtt      120
ggggggtaag tccggcacacg gcaccccttac acagatgaca accagggaaac tgaagtccac      180
ctggagatgg ggcaggatgg cacgggtata gggaccacac accgcagcccc agaaagtctc      240
ctggagatca aagcctgaa gccagaggc atccctgtc tgggtgtcaa ggcctccagg      300
tttcttgcc aacaaccaga cggaaaccctg tatggatcgc ctcacttgc tcctgaggcc      360
tgcagttca gggagcttt gcttgaggat ggataacaatg tggatccaatc tgaagtccat      420
ggcctgcccc tgcgcctgcc ccagagggac tctccaaacc aggccccagc atcctgggg      480
cctgtgcccc ccctgcccagt gccaggactg ctccaccaggc cccaggagct accagggttc      540

```

US 9,475,856 B2

329**330**

-continued

ctggccccag aacctccaga tgtggctcc tctgaccac tgagcatggt gggaccttg	600
cagggccgaa gccccagcta tgcttcctga	630

<210> SEQ ID NO 200

<211> LENGTH: 579

<212> TYPE: DNA

<213> ORGANISM: Ovis aries

<400> SEQUENCE: 200

atgggctgg acgaggccaa gttcaagcac ttggactgt gggccctgt gctggctgtc	60
ctcctgttag gAACCTGCCG ggcgcatacca attccagact ccagccccct cctccagttt	120
ggggggccaag tccggcagcg gtacctctac acggatgtat cccaggagac agaggcccac	180
ctggagatca gggccgatgg cacagtggtgg gggccggccc gccagagtcc cgaaagtctc	240
ttggagctga aagccctgaa gccaggagtc attcagatct ttggagttaa aacatccagg	300
ttcctgtgcc agggggccaga tgggaagctg tatggatcgc tgcactttga ccccaaagcc	360
tgcagcttcc gggagctgct tcttgaagat gggtaataatg tctaccatgc ggagacctg	420
ggcccttccac tccgcctgcc gccgcagcgc tcatccaacc gggacccggc cccgcgggga	480
cctccgaagc cccagctaca ctttttgaag acgtccgctg tgcagttactg gccacgttat	540
gagaaggctcc cagttttctc gcaccccttc cccggctga	579

<210> SEQ ID NO 201

<211> LENGTH: 405

<212> TYPE: DNA

<213> ORGANISM: Pan paniscus

<400> SEQUENCE: 201

atggactcgg acgagacccgg gttcgagcac tcaggactgt gggtttctgt gctggctgg	60
cttctgtgg gagcctgccca ggcacacccc atccctgact ccagtcctct cctgcaattc	120
ggggggccaag tccggcagcg gtacctctac acagatgtat cccagcagac agaagcccac	180
ctggagatca gggaggatgg gacgggtgggg ggcgcgtctg accagagccc cgaaagtctc	240
ctgcagctga aagccctgaa gccgggagtt attcaaatct ttggagttcaa gacatccagg	300
ttcctgtgcc agaggccaga tggggccctg tatggatcgg tgagtttcca ggaccctcct	360
caccacccac catgctcctc ctatatgtcg ccctcacagc ctggg	405

<210> SEQ ID NO 202

<211> LENGTH: 351

<212> TYPE: DNA

<213> ORGANISM: Macaca fascicularis

<400> SEQUENCE: 202

atggatagcg atgaaaacccgg ctttgaacat agcggcctgt gggtgcgggt gctggccggc	60
ctgctgtgg ggcgcgtgccca ggcgcataccg attccggata gcaagcccgct gctgcagttt	120
ggcggccagg tgcgcagcg ctatctgtat accgatgtat cgcagcagac cgaagcgcac	180
ctggaaattc gcgaagatgg caccgtgggc ggcgcggcgc atcagagccc ggaaagctg	240
ctgcagctga aagcgctgaa accgggcgtg attcagattc ttggcgtgaa aaccagccgc	300
tttctgtgcc agaaaacccgg tggcgcgctg tatggcagcgc tgagctttta a	351

<210> SEQ ID NO 203

<211> LENGTH: 314

<212> TYPE: DNA

<213> ORGANISM: Mesocricetus auratus

-continued

<400> SEQUENCE: 203

```

ggtcatcaa atcctgggtg tcaaggctgc taggtttcct tgccagcaac cagacggaag      60
cctgtacgga tcgcctcaact tcgatcccga ggcctgcagt ttccgggagc tcctgcttga     120
ggatggatac aatgtgtacc agtcggaagc ccacggcctg cccctgcgcc tgccccagag     180
ggacgctccg agccagcccc cagcatcctg gggaccggtg cgcttctgc cagtgcccg     240
actgttccag ccgccccacg acctcccagg ggcggccggc ccagagcctc cggacgtggg     300
ctcctccgac ccac                                         314

```

<210> SEQ ID NO 204

<211> LENGTH: 456

<212> TYPE: DNA

<213> ORGANISM: Nile tilapia

<400> SEQUENCE: 204

```

atgtatttgc agatgaacat gnatggaga gtcacaggaa gtgatgtca gacacattac      60
agttttagtc agctgaaatc agttaaacca ggccatgtaa tcattaaagg accatcatca    120
tctcttttc tctgtgtgga cagcgaaggc aatctgagag ggcagagtca ctactcagaa    180
accagctgca ctttcagaga aatgctgctg gctgacggat acacccgtt catttctca    240
caatatggat ttccatgtc actggcatca agacattccc cagatcgaca cgcgttccc    300
tttacgggt tcctaccact gaggaataac ttgaaaacgg atagcgtatc agagcagctg    360
ccaaacaatc agagactctt caacgtggac tctgatgacc ttcttggaat gggctaaat    420
tctatggca gtcctcagtt ttctatggac aaataa                                         456

```

<210> SEQ ID NO 205

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 205

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val
1				5				10			15				

Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His
	20				25					30					

Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser
	35				40				45						

Pro	Glu	Ser	Leu	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln
50				55			60								

Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Gln	Arg	Pro	Asp	Gly
65				70				75			80				

Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg
	85				90				95						

Glu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His
	100				105			110						

Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro
	115				120				125						

Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Gly	Leu	Pro	Pro		
130				135			140								

Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro	Gln	Pro	Pro	Asp	Val
145				150			155			160					

-continued

Gly Ser Ser Asp Pro Leu Ser Met Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 206

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 206

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145 150 155 160

Gly Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 207

<211> LENGTH: 187

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 207

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

-continued

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Leu Pro Met Val Pro
 130 135 140

Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser
 145 150 155 160

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
 165 170 175

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 180 185

<210> SEQ ID NO 208

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 208

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gln Val
 1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
 145 150 155 160

Gly Ser Ser Asp Pro Leu Ser Met Val Thr Gly Leu Glu Ala Val Arg
 165 170 175

Ser Pro Ser Phe Glu Lys
 180

<210> SEQ ID NO 209

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 209

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gln Val
 1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His

-continued

20	25	30
Leu	Glu	Ile
Arg	Glu	Asp
Gly	Thr	Val
Gly	Gly	Ala
Ala	Asp	Gln
Ser		Ser
35	40	45
Pro	Glu	Ser
Leu	Leu	Gln
Lys	Ala	Leu
50	55	60
Ile	Leu	Gly
Val	Lys	Thr
Ser	Arg	Phe
Leu	Cys	Met
65	70	75
Ala	Tyr	Gly
Ser	Leu	His
Phe	Asp	Pro
Glu	Ala	Cys
85	90	95
Glu	Leu	Leu
Glu	Asp	Gly
Tyr	Asn	Val
100	105	110
Gly	Leu	Pro
Pro	Leu	His
Gly	Asn	Lys
Ser	Pro	His
115	120	125
Ala	Pro	Arg
Gly	Pro	Ala
Arg	Leu	Phe
130	135	140
Ala	Leu	Pro
Glu	Pro	Pro
Pro	Gly	Ile
145	150	155
Gly	Ser	Met
Asp	Pro	Phe
Gly	Leu	Val
165	170	175
Ser	Pro	Ser
Phe	Glu	Lys
180		

<210> SEQ ID NO 210

<211> LENGTH: 187

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 210

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	
1								5	10	15						
Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	Val	Arg
									20	25	30					
Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser	
									35	40	45					
Pro	Glu	Ser	Leu	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	
									50	55	60					
Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Met	Arg	Pro	Asp	Gly	
65									70	75	80					
Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg	
									85	90	95					
Glu	Leu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	
									100	105	110					
Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro	
									115	120	125					
Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Leu	Leu	Pro	Met	Val	Pro	
									130	135	140					
Glu	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu	Ser	Asp	Met	Phe	Ser	
145									150	155	160					
Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr	Gly	
									165	170	175					
Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys						
									180	185						

-continued

<210> SEQ ID NO 211
<211> LENGTH: 546
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 211

caccccatcc	ctgactccag	tcctctcctg	caattcgaaa	gcacaagtccg	gcageggtag	60
ctctacacag	atgatgccc	gcagacagaa	gcccacctgg	agatcaggga	ggatgggacg	120
gtggggggcg	ctgctgacca	gagccccgaa	agtctcctgc	agctgaaagc	cttgaagccg	180
ggagttattc	aatcttggg	agtcaagaca	tccaggttcc	tgtgccagcg	gccagatggg	240
gccctgtatg	gatcgctcca	cattgaccct	gaggcctgca	gcttccggga	gctgttttt	300
gaggacggat	acaatgttta	ccagtcgaaa	gcccacggcc	tcccgctgca	cctgccaggg	360
aacaagtccc	cacaccggaa	ccctgcaccc	cgaggaccag	ctcgcttcc	gccactacca	420
ggcctgcccc	ccgcactccc	ggagccaccc	ggaatctgg	ccccccagcc	ccccgatgtg	480
ggctcctcgg	accctctgag	catggtgggaa	ctggaggccg	tgaggagttcc	cagctttag	540
aagtaa						546

<210> SEQ ID NO 212
<211> LENGTH: 549
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 212

caccccatcc	ctgactccag	tcctctcctg	caattcgaaa	gcacaagtccg	gcageggtag	60
ctctacacag	atgatgccc	gcagacagaa	gcccacctgg	agatcaggga	ggatgggacg	120
gtggggggcg	ctgctgacca	gagccccgaa	agtctcctgc	agctgaaagc	cttgaagccg	180
ggagttattc	aatcttggg	agtcaagaca	tccaggttcc	tgtgccagcg	gccagatggg	240
gccctgtatg	gatcgctcca	cattgaccct	gaggcctgca	gcttccggga	gctgttttt	300
gaggacggat	acaatgttta	ccagtcgaaa	gcccacggcc	tcccgctgca	cctgecaggg	360
aacaagtccc	cacaccggaa	ccctgcaccc	cgaggaccag	ctcgcttcc	gccactacca	420
ggcctgcccc	ccgcactccc	ggagccaccc	ggaatctgg	ccccccagcc	ccccgatgtg	480
ggctccatgg	accctttgg	gcttgtcacc	ggactggagg	ccgtgaggag	tcccagctt	540
gagaagtaa						549

<210> SEQ ID NO 213
<211> LENGTH: 564
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 213

caccccatcc	ctgactccag	tcctctcctg	caattcgaaa	gcacaagtccg	gcagcggtac	60
ctctacacag	atgatgccc	gcagacagaa	gcccacctgg	agatcaggga	ggatgggacg	120
gtggggggcg	ctgctgacca	gagccccgaa	agtctcctgc	agctgaaagc	cttgaagccg	180
ggagttattc	aatcttggg	agtcaagaca	tccaggttcc	tgtgccagcg	gccagatggg	240

-continued

gccctgtatg gatcgctcca ctttgaccct gaggcctgca gttccggga gctgcttctt	300
gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg	360
aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctgcgttcct gccactactg	420
cccatggtcc cagaggagcc tgaggaccc agggggccact tggaatctga catgttctct	480
tcgccccctgg agaccgacag catggaccctt tttgggcttg tcacccgact ggaggccgtg	540
aggagtccta gctttgagaa gtaa	564

<210> SEQ ID NO 214
<211> LENGTH: 546
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 214

caccccatcc ctgactccag tcctctcctg caattcgggg gccaagtccg gcagcggtac	60
ctctacacag atgatgccc gcagacagaa gcccacctgg agatcaggga ggatgggacg	120
gtggggggcg ctgctgacca gagcccccggaa agtctcctgc agctgaaagc cttgaagccg	180
ggagttattc aaatcttggg agtcaagaca tccaggttcc tgcgtccatg gccagatggg	240
gccctgtatg gatcgctcca ctttgaccct gaggcctgca gttccggga gctgcttctt	300
gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg	360
aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctgcgttcct gccactacca	420
ggcctgcccccc cgcaactccc ggagccaccc ggaatcctgg ccccccagcc ccccgatgtg	480
ggctccctcg acccctctgag catggtggga ctggaggccg tgaggagttcc cagctttag	540
aagtaa	546

<210> SEQ ID NO 215
<211> LENGTH: 549
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 215

caccccatcc ctgactccag tcctctcctg caattcgggg gccaagtccg gcagcggtac	60
ctctacacag atgatgccc gcagacagaa gcccacctgg agatcaggga ggatgggacg	120
gtggggggcg ctgctgacca gagcccccggaa agtctcctgc agctgaaagc cttgaagccg	180
ggagttattc aaatcttggg agtcaagaca tccaggttcc tgcgtccatg gccagatggg	240
gccctgtatg gatcgctcca ctttgaccct gaggcctgca gttccggga gctgcttctt	300
gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg	360
aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctgcgttcct gccactacca	420
ggcctgcccccc cgcaactccc ggagccaccc ggaatcctgg ccccccagcc ccccgatgtg	480
ggctccatgg acccatttgg gcttgtcacc ggactggagg ccgtgaggag tcccgacttt	540
gagaagtaa	549

<210> SEQ ID NO 216
<211> LENGTH: 564
<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Nucleic Acid Molecule Encoding a Chimeric Protein

<400> SEQUENCE: 216

```
caccccatcc ctgactccag tccttcctcg caattcgggg gccaagtccg gcagcggtac      60
ctctacacag atgatgccca gcagacagaa gcccacctgg agatcaggga ggatgggacg     120
gtggggggcg ctgctgacca gagcccccga agtctcctgc agctgaaagc cttgaagccg     180
ggagttattc aaatcttggg agtcaagaca tccaggttcc tgcataatg gccagatggg     240
gcctgtatg gatcgctcca ctttgaccct gaggcctgca gttccggga gctgttctt     300
gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgcccagg     360
aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctgcgttcc gccactactg     420
cccatggtcc cagaggagcc tgaggaccc aggggccact tgaaatctga catgttctt     480
tcgccccctgg agaccgacag catggaccca ttggggcttg tcaccggact ggaggccgtg     540
aggagtccca gctttgagaa gtaa                                         564
```

<210> SEQ ID NO 217

<211> LENGTH: 1044

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 217

Met	Lys	Pro	Gly	Cys	Ala	Ala	Gly	Ser	Pro	Gly	Asn	Glu	Trp	Ile	Phe
1															
					5			10				15			

Phe	Ser	Thr	Asp	Glu	Ile	Thr	Thr	Arg	Tyr	Arg	Asn	Thr	Met	Ser	Asn
					20			25				30			

Gly	Gly	Leu	Gln	Arg	Ser	Val	Ile	Leu	Ser	Ala	Leu	Ile	Leu	Leu	Arg
		35				40					45				

Ala	Val	Thr	Gly	Phe	Ser	Gly	Asp	Gly	Arg	Ala	Ile	Trp	Ser	Lys	Asn
	50				55						60				

Pro	Asn	Phe	Thr	Pro	Val	Asn	Glu	Ser	Gln	Leu	Phe	Leu	Tyr	Asp	Thr
	65				70				75		80				

Phe	Pro	Lys	Asn	Phe	Phe	Trp	Gly	Ile	Gly	Thr	Gly	Ala	Leu	Gln	Val
	85					90				95					

Glu	Gly	Ser	Trp	Lys	Lys	Asp	Gly	Lys	Pro	Ser	Ile	Trp	Asp	His
	100					105					110			

Phe	Ile	His	Thr	His	Leu	Lys	Asn	Val	Ser	Ser	Thr	Asn	Gly	Ser	Ser
	115				120				125						

Asp	Ser	Tyr	Ile	Phe	Leu	Glu	Lys	Asp	Leu	Ser	Ala	Leu	Asp	Phe	Ile
	130				135				140						

Gly	Val	Ser	Phe	Tyr	Gln	Phe	Ser	Ile	Ser	Trp	Pro	Arg	Leu	Phe	Pro
	145				150			155		160					

Asp	Gly	Ile	Val	Thr	Val	Ala	Asn	Ala	Lys	Gly	Leu	Gln	Tyr	Tyr	Ser
	165				170				175						

Thr	Leu	Leu	Asp	Ala	Leu	Val	Leu	Arg	Asn	Ile	Glu	Pro	Ile	Val	Thr
	180				185				190						

Leu	Tyr	His	Trp	Asp	Leu	Pro	Leu	Ala	Leu	Gln	Glu	Lys	Tyr	Gly	Gly
	195				200				205						

Trp	Lys	Asn	Asp	Thr	Ile	Ile	Asp	Ile	Phe	Asn	Asp	Tyr	Ala	Thr	Tyr
	210				215			220							

Cys	Phe	Gln	Met	Phe	Gly	Asp	Arg	Val	Lys	Tyr	Trp	Ile	Thr	Ile	His
	225				230			235		240					

-continued

Asn Pro Tyr Leu Val Ala Trp His Gly Tyr Gly Thr Gly Met His Ala
 245 250 255
 Pro Gly Glu Lys Gly Asn Leu Ala Ala Val Tyr Thr Val Gly His Asn
 260 265 270
 Leu Ile Lys Ala His Ser Lys Val Trp His Asn Tyr Asn Thr His Phe
 275 280 285
 Arg Pro His Gln Lys Gly Trp Leu Ser Ile Thr Leu Gly Ser His Trp
 290 295 300
 Ile Glu Pro Asn Arg Ser Glu Asn Thr Met Asp Ile Phe Lys Cys Gln
 305 310 315 320
 Gln Ser Met Val Ser Val Leu Gly Trp Phe Ala Asn Pro Ile His Gly
 325 330 335
 Asp Gly Asp Tyr Pro Glu Gly Met Arg Lys Lys Leu Phe Ser Val Leu
 340 345 350
 Pro Ile Phe Ser Glu Ala Glu Lys His Glu Met Arg Gly Thr Ala Asp
 355 360 365
 Phe Phe Ala Phe Ser Phe Gly Pro Asn Asn Phe Lys Pro Leu Asn Thr
 370 375 380
 Met Ala Lys Met Gly Gln Asn Val Ser Leu Asn Leu Arg Glu Ala Leu
 385 390 395 400
 Asn Trp Ile Lys Leu Glu Tyr Asn Asn Pro Arg Ile Leu Ile Ala Glu
 405 410 415
 Asn Gly Trp Phe Thr Asp Ser Arg Val Lys Thr Glu Asp Thr Thr Ala
 420 425 430
 Ile Tyr Met Met Lys Asn Phe Leu Ser Gln Val Leu Gln Ala Ile Arg
 435 440 445
 Leu Asp Glu Ile Arg Val Phe Gly Tyr Thr Ala Trp Ser Leu Leu Asp
 450 455 460
 Gly Phe Glu Trp Gln Asp Ala Tyr Thr Ile Arg Arg Gly Leu Phe Tyr
 465 470 475 480
 Val Asp Phe Asn Ser Lys Gln Lys Glu Arg Lys Pro Lys Ser Ser Ala
 485 490 495
 His Tyr Tyr Lys Gln Ile Ile Arg Glu Asn Gly Phe Ser Leu Lys Glu
 500 505 510
 Ser Thr Pro Asp Val Gln Gly Gln Phe Pro Cys Asp Phe Ser Trp Gly
 515 520 525
 Val Thr Glu Ser Val Leu Lys Pro Glu Ser Val Ala Ser Ser Pro Gln
 530 535 540
 Phe Ser Asp Pro His Leu Tyr Val Trp Asn Ala Thr Gly Asn Arg Leu
 545 550 555 560
 Leu His Arg Val Glu Gly Val Arg Leu Lys Thr Arg Pro Ala Gln Cys
 565 570 575
 Thr Asp Phe Val Asn Ile Lys Lys Gln Leu Glu Met Leu Ala Arg Met
 580 585 590
 Lys Val Thr His Tyr Arg Phe Ala Leu Asp Trp Ala Ser Val Leu Pro
 595 600 605
 Thr Gly Asn Leu Ser Ala Val Asn Arg Gln Ala Leu Arg Tyr Tyr Arg
 610 615 620
 Cys Val Val Ser Glu Gly Leu Lys Leu Gly Ile Ser Ala Met Val Thr
 625 630 635 640
 Leu Tyr Tyr Pro Thr His Ala His Leu Gly Leu Pro Glu Pro Leu Leu
 645 650 655

-continued

His Ala Asp Gly Trp Leu Asn Pro Ser Thr Ala Glu Ala Phe Gln Ala
 660 665 670
 Tyr Ala Gly Leu Cys Phe Gln Glu Leu Gly Asp Leu Val Lys Leu Trp
 675 680 685
 Ile Thr Ile Asn Glu Pro Asn Arg Leu Ser Asp Ile Tyr Asn Arg Ser
 690 695 700
 Gly Asn Asp Thr Tyr Gly Ala Ala His Asn Leu Leu Val Ala His Ala
 705 710 715 720
 Leu Ala Trp Arg Leu Tyr Asp Arg Gln Phe Arg Pro Ser Gln Arg Gly
 725 730 735
 Ala Val Ser Leu Ser Leu His Ala Asp Trp Ala Glu Pro Ala Asn Pro
 740 745 750
 Tyr Ala Asp Ser His Trp Arg Ala Ala Glu Arg Phe Leu Gln Phe Glu
 755 760 765
 Ile Ala Trp Phe Ala Glu Pro Leu Phe Lys Thr Gly Asp Tyr Pro Ala
 770 775 780
 Ala Met Arg Glu Tyr Ile Ala Ser Lys His Arg Arg Gly Leu Ser Ser
 785 790 795 800
 Ser Ala Leu Pro Arg Leu Thr Glu Ala Glu Arg Arg Leu Leu Lys Gly
 805 810 815
 Thr Val Asp Phe Cys Ala Leu Asn His Phe Thr Thr Arg Phe Val Met
 820 825 830
 His Glu Gln Leu Ala Gly Ser Arg Tyr Asp Ser Asp Arg Asp Ile Gln
 835 840 845
 Phe Leu Gln Asp Ile Thr Arg Leu Ser Ser Pro Thr Arg Leu Ala Val
 850 855 860
 Ile Pro Trp Gly Val Arg Lys Leu Leu Arg Trp Val Arg Arg Asn Tyr
 865 870 875 880
 Gly Asp Met Asp Ile Tyr Ile Thr Ala Ser Gly Ile Asp Asp Gln Ala
 885 890 895
 Leu Glu Asp Asp Arg Leu Arg Lys Tyr Tyr Leu Gly Lys Tyr Leu Gln
 900 905 910
 Glu Val Leu Lys Ala Tyr Leu Ile Asp Lys Val Arg Ile Lys Gly Tyr
 915 920 925
 Tyr Ala Phe Lys Leu Ala Glu Glu Lys Ser Lys Pro Arg Phe Gly Phe
 930 935 940
 Phe Thr Ser Asp Phe Lys Ala Lys Ser Ser Ile Gln Phe Tyr Asn Lys
 945 950 955 960
 Val Ile Ser Ser Arg Gly Phe Pro Phe Glu Asn Ser Ser Arg Cys
 965 970 975
 Ser Gln Thr Gln Glu Asn Thr Glu Cys Thr Val Cys Leu Phe Leu Val
 980 985 990
 Gln Lys Lys Pro Leu Ile Phe Leu Gly Cys Cys Phe Phe Ser Thr Leu
 995 1000 1005
 Val Leu Leu Leu Ser Ile Ala Ile Phe Gln Arg Gln Lys Arg Arg
 1010 1015 1020
 Lys Phe Trp Lys Ala Lys Asn Leu Gln His Ile Pro Leu Lys Lys
 1025 1030 1035
 Gly Lys Arg Val Val Ser
 1040

<210> SEQ ID NO 218
 <211> LENGTH: 1043
 <212> TYPE: PRT

-continued

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 218

Met	Lys	Thr	Gly	Cys	Ala	Ala	Gly	Ser	Pro	Gly	Asn	Glu	Trp	Ile	Phe
1				5				10				15			

Phe	Ser	Ser	Asp	Glu	Arg	Asn	Thr	Arg	Ser	Arg	Lys	Thr	Met	Ser	Asn
20				25				30							

Arg	Ala	Leu	Gln	Arg	Ser	Ala	Val	Leu	Ser	Ala	Phe	Val	Leu	Leu	Arg
35				40				45							

Ala	Val	Thr	Gly	Phe	Ser	Gly	Asp	Gly	Lys	Ala	Ile	Trp	Asp	Lys	Lys
50				55			60								

Gln	Tyr	Val	Ser	Pro	Val	Asn	Pro	Ser	Gln	Leu	Phe	Leu	Tyr	Asp	Thr
65				70			75		80						

Phe	Pro	Lys	Asn	Phe	Ser	Trp	Gly	Val	Gly	Thr	Gly	Ala	Phe	Gln	Val
85				90			95								

Glu	Gly	Ser	Trp	Lys	Thr	Asp	Gly	Arg	Gly	Pro	Ser	Ile	Trp	Asp	Arg
100				105			110								

Tyr	Val	Tyr	Ser	His	Leu	Arg	Gly	Val	Asn	Gly	Thr	Asp	Arg	Ser	Thr
115				120			125								

Asp	Ser	Tyr	Ile	Phe	Leu	Glu	Lys	Asp	Leu	Leu	Ala	Leu	Asp	Phe	Leu
130				135			140								

Gly	Val	Ser	Phe	Tyr	Gln	Phe	Ser	Ile	Ser	Trp	Pro	Arg	Leu	Phe	Pro
145				150			155			160					

Asn	Gly	Thr	Val	Ala	Ala	Val	Asn	Ala	Gln	Gly	Leu	Arg	Tyr	Tyr	Arg
165				170			175								

Ala	Leu	Leu	Asp	Ser	Leu	Val	Leu	Arg	Asn	Ile	Glu	Pro	Ile	Val	Thr
180				185			190								

Leu	Tyr	His	Trp	Asp	Leu	Pro	Leu	Thr	Leu	Gln	Glu	Glu	Tyr	Gly	Gly
195				200			205								

Trp	Lys	Asn	Ala	Thr	Met	Ile	Asp	Leu	Phe	Asn	Asp	Tyr	Ala	Thr	Tyr
210				215			220								

Cys	Phe	Gln	Thr	Phe	Gly	Asp	Arg	Val	Lys	Tyr	Trp	Ile	Thr	Ile	His
225				230			235			240					

Asn	Pro	Tyr	Leu	Val	Ala	Trp	His	Gly	Phe	Gly	Thr	Gly	Met	His	Ala
245				250			255								

Pro	Gly	Glu	Lys	Gly	Asn	Leu	Thr	Ala	Val	Tyr	Thr	Val	Gly	His	Asn
260				265			270								

Leu	Ile	Lys	Ala	His	Ser	Lys	Val	Trp	His	Asn	Tyr	Asp	Lys	Asn	Phe
275				280			285								

Arg	Pro	His	Gln	Lys	Gly	Trp	Leu	Ser	Ile	Thr	Leu	Gly	Ser	His	Trp
290				295			300								

Ile	Glu	Pro	Asn	Arg	Thr	Asp	Asn	Met	Glu	Asp	Val	Ile	Asn	Cys	Gln
305				310			315			320					

His	Ser	Met	Ser	Ser	Val	Leu	Gly	Trp	Phe	Ala	Asn	Pro	Ile	His	Gly
325				330			335								

Asp	Gly	Asp	Tyr	Pro	Glu	Phe	Met	Lys	Thr	Gly	Ala	Met	Ile	Pro	Glu
340				345			350								

Phe	Ser	Glu	Ala	Glu	Lys	Glu	Glu	Val	Arg	Gly	Thr	Ala	Asp	Phe	Phe
355				360			365								

Ala	Phe	Ser	Phe	Gly	Pro	Asn	Asn	Phe	Arg	Pro	Ser	Asn	Thr	Val	Val
370				375			380								

Lys	Met	Gly	Gln	Asn	Val	Ser	Leu	Asn	Leu	Arg	Gln	Val	Leu	Asn	Trp
385				390			395			400					

-continued

Ile Lys Leu Glu Tyr Asp Asp Pro Gln Ile Leu Ile Ser Glu Asn Gly
405 410 415

Trp Phe Thr Asp Ser Tyr Ile Lys Thr Glu Asp Thr Thr Ala Ile Tyr
420 425 430

Met Met Lys Asn Phe Leu Asn Gln Val Leu Gln Ala Ile Lys Phe Asp
435 440 445

Glu Ile Arg Val Phe Gly Tyr Thr Ala Trp Thr Leu Leu Asp Gly Phe
450 455 460

Glu Trp Gln Asp Ala Tyr Thr Arg Arg Gly Leu Phe Tyr Val Asp
465 470 475 480

Phe Asn Ser Glu Gln Lys Glu Arg Lys Pro Lys Ser Ser Ala His Tyr
485 490 495

Tyr Lys Gln Ile Ile Gln Asp Asn Gly Phe Pro Leu Lys Glu Ser Thr
500 505 510

Pro Asp Met Lys Gly Arg Phe Pro Cys Asp Phe Ser Trp Gly Val Thr
515 520 525

Glu Ser Val Leu Lys Pro Glu Phe Thr Val Ser Ser Pro Gln Phe Thr
530 535 540

Asp Pro His Leu Tyr Val Trp Asn Val Thr Gly Asn Arg Leu Leu Tyr
545 550 555 560

Arg Val Glu Gly Val Arg Leu Lys Thr Arg Pro Ser Gln Cys Thr Asp
565 570 575

Tyr Val Ser Ile Lys Lys Arg Val Glu Met Leu Ala Lys Met Lys Val
580 585 590

Thr His Tyr Gln Phe Ala Leu Asp Trp Thr Ser Ile Leu Pro Thr Gly
595 600 605

Asn Leu Ser Lys Val Asn Arg Gln Val Leu Arg Tyr Tyr Arg Cys Val
610 615 620

Val Ser Glu Gly Leu Lys Leu Gly Val Phe Pro Met Val Thr Leu Tyr
625 630 635 640

His Pro Thr His Ser His Leu Gly Leu Pro Leu Pro Leu Ser Ser
645 650 655

Gly Gly Trp Leu Asn Met Asn Thr Ala Lys Ala Phe Gln Asp Tyr Ala
660 665 670

Glu Leu Cys Phe Arg Glu Leu Gly Asp Leu Val Lys Leu Trp Ile Thr
675 680 685

Ile Asn Glu Pro Asn Arg Leu Ser Asp Met Tyr Asn Arg Thr Ser Asn
690 695 700

Asp Thr Tyr Arg Ala Ala His Asn Leu Met Ile Ala His Ala Gln Val
705 710 715 720

Trp His Leu Tyr Asp Arg Gln Tyr Arg Pro Val Gln His Gly Ala Val
725 730 735

Ser Leu Ser Leu His Cys Asp Trp Ala Glu Pro Ala Asn Pro Phe Val
740 745 750

Asp Ser His Trp Lys Ala Ala Glu Arg Phe Leu Gln Phe Glu Ile Ala
755 760 765

Trp Phe Ala Asp Pro Leu Phe Lys Thr Gly Asp Tyr Pro Ser Val Met
770 775 780

Lys Glu Tyr Ile Ala Ser Lys Asn Gln Arg Gly Leu Ser Ser Ser Val
785 790 795 800

Leu Pro Arg Phe Thr Ala Lys Glu Ser Arg Leu Val Lys Gly Thr Val
805 810 815

Asp Phe Tyr Ala Leu Asn His Phe Thr Thr Arg Phe Val Ile His Lys

US 9,475,856 B2

353

354

-continued

820	825	830
Gln Leu Asn Thr Asn Arg Ser Val Ala Asp Arg Asp Val Gln Phe Leu		
835	840	845
Gln Asp Ile Thr Arg Leu Ser Ser Pro Ser Arg Leu Ala Val Thr Pro		
850	855	860
Trp Gly Val Arg Lys Leu Leu Ala Trp Ile Arg Arg Asn Tyr Arg Asp		
865	870	875
Arg Asp Ile Tyr Ile Thr Ala Asn Gly Ile Asp Asp Leu Ala Leu Glu		
885	890	895
Asp Asp Gln Ile Arg Lys Tyr Tyr Leu Glu Lys Tyr Val Gln Glu Ala		
900	905	910
Leu Lys Ala Tyr Leu Ile Asp Lys Val Lys Ile Lys Gly Tyr Tyr Ala		
915	920	925
Phe Lys Leu Thr Glu Glu Lys Ser Lys Pro Arg Phe Gly Phe Phe Thr		
930	935	940
Ser Asp Phe Arg Ala Lys Ser Ser Val Gln Phe Tyr Ser Lys Leu Ile		
945	950	955
Ser Ser Ser Gly Leu Pro Ala Glu Asn Arg Ser Pro Ala Cys Gly Gln		
965	970	975
Pro Ala Glu Asp Thr Asp Cys Thr Ile Cys Ser Phe Leu Val Glu Lys		
980	985	990
Lys Pro Leu Ile Phe Phe Gly Cys Cys Phe Ile Ser Thr Leu Ala Val		
995	1000	1005
Leu Leu Ser Ile Thr Val Phe His His Gln Lys Arg Arg Lys Phe		
1010	1015	1020
Gln Lys Ala Arg Asn Leu Gln Asn Ile Pro Leu Lys Lys Gly His		
1025	1030	1035
Ser Arg Val Phe Ser		
1040		

<210> SEQ ID NO 219

<211> LENGTH: 3135

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 219

atgaagccag gctgtgcggc aggatctcca ggaaatgaat ggattttctt cagcactgat	60
gaaaataacca cacgtatag gaatacaatg tccaacgggg gattgcaaag atctgtcatc	120
ctgtcagcac ttattctgct acgagctgtt actggattct ctggagatgg aagagctata	180
tggtctaaaa atcctaattt tactccgta aatgaaagtcc agctgtttct ctagacact	240
ttccctaaaa acttttctg gggatttggg actggagcat tgcaagtgga agggagttgg	300
aagaaggatg gaaaaggacc ttcttatatgg gatcatttca tccacacaca ccttaaaaat	360
gtcagcagca cgaatggttc cagtacatgat tatatttttc tggaaaaaga cttatcagcc	420
ctggatttta taggagtttcc ttttatcaa tttcaattt cctggccaag gctttcccc	480
gatggatag taacagttgc caacgcaaaa ggtctgcagt actacagttac tcttctggac	540
gctctagtgc ttagaaacat tgaacctata gttactttat accactggga tttgccttg	600
gcactacaag aaaaatatgg ggggtggaaa aatgataccca taatagatat cttcaatgac	660
tatgccacat actgtttcca gatgtttggg gaccgtgtca aatattggat tacaattcac	720
aacccatatac tagtggcttg gcatgggtat gggacaggtta tgcgtgccccc tggagagaag	780
ggaaatttag cagctgtcta cactgtggga cacaacttga tcaaggctca ctcgaaagtt	840

-continued

tggcataact acaacacaca tttccgcccc aatcagaagg gttggttatc gatcacgtt	900
ggatctcatt ggatcgagcc aaaccggctcg gaaaacacga tggatattatt caaatgtcaa	960
caatccatgg tttctgtgct tggatggttt gccaacccta tccatgggga tggcgactat	1020
ccagagggga tgagaaagaa gttgttctcc gttctaccca ttttctctga agcagagaag	1080
catgagatga gaggcacagc tgatttcttt gccttttctt ttggacccaa caacttcaag	1140
cccctaaaca ccatggctaa aatgggacaa aatgtttcac ttaatttaag agaagcgctg	1200
aactggatta aactggaata caadaaccct cgaatcttga ttgctgagaa tggctggtc	1260
acagacagtc gtgtaaaaac agaagacacc acggccatct acatgtgaa gaatttcctc	1320
agccaggtgc ttcaagcaat aaggttagat gaaatacggag tgtttggtt tactgcctgg	1380
tctctcttgg atggcttga atggcaggat gcttacacca tccgcccagg attattttat	1440
gtggatttta acagtaaaca gaaagagcgg aaacctaagt ctccagcaca ctactacaaa	1500
cagatcatac gagaaaaatgg ttttcttta aaagagtccca cgccagatgt gcagggccag	1560
tttccctgtg acttctctg ggggtgtcact gaatctgttc ttaagccoga gtctgtggct	1620
tegtccccac agttcagcga tcctcatctg tacgtgtggaa acggccactgg caacagactg	1680
ttgcacccag tggaaagggtt gaggctgaaa acacgacccg ctcaatgcac agatttgtt	1740
aacatcaaaa aacaacttga gatgttggca agaatgaaag tcacccacta ccggtttgct	1800
ctggatttggg ctcggctct tcctactggc aacctgtccg cggtgaaccc acaggccctg	1860
aggtactaca ggtgegttgtt cagtggggg ctgaagcttgc gcatctccgc gatggtcacc	1920
ctgttattatc cgacccacgc ccaccttaggc ctccccgagc ctctgttgca tgccgacggg	1980
tggctgaacc catcgacggc cgaggcccttc caggcctacg ctgggtgtg cttccaggag	2040
ctgggggacc tggtaagct ctggatcacc atcaacgagc ctaacccgct aagtgacatc	2100
tacaaccgc ctggcaacga cacctacggg gggcgacaca acctgttgtt ggcccacgcc	2160
ctggcctggc gcctctacga cggcagttc agggccctcac agcgcggggc cgtgtcgctg	2220
tcgctgcacg cggactgggc ggaacccgc aacccctatg ctgactcgca ctggaggggc	2280
gccgagcgct tcctgcagtt cgagatcgcc tggttcgccg agccgcetttt caagaccggg	2340
gactaccggc cggccatgag ggaatacatt gctccaagg accgaegggg gctttccagc	2400
tcggccctgc cgcgectcac cgaggccgaa aggaggctgc tcaagggcac ggtcgacttc	2460
tcgcgcgtca accacttac cactagggtc gtgtatgcacg agcagctggc cggcagccgc	2520
tacgactcg acagggacat ccagtttctg caggacatca cccgcctgag ctccccacgc	2580
cgccctggctg tgattccctg gggggtgcgc aagctgtgc ggtgggtccg gaggaactac	2640
ggcgacatgg acatttacat caccggccagt ggcacatgcacg accaggctct ggaggatgac	2700
cggtctccgg agtactacat agggaaagtac cttcaggagg tgctgaaagc atacctgatt	2760
gataaaagtca gaatcaaagg ctattatgca ttcaaactgg ctgaagagaa atctaaaccc	2820
agatttggat tcttcacatc tgattttaaa gctaaatctt caatacaattt ttacaacaaa	2880
gtgtatcgca gcagggggctt ccctttttagt aacagtagttt ctagatgcag tcagacccaa	2940
aaaaatacag agtgcactgt ctgttatttc ctgtgcaga agaaaaccact gatattctg	3000
ggttgttgct tcttccttccac cctgggttcta ctcttatcaa ttgcccatttt tcaaaggcag	3060
aagagaagaa agttttggaa agcaaaaaac ttacaacaca taccattaaa gaaaggcaag	3120
agagttgtta gctaa	3135

-continued

<210> SEQ ID NO 220
<211> LENGTH: 3132
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 220

atgaagacag gctgtcgac	agggtctccg	ggaaatgaat ggattttctt cagctctgat	60
gaaaagaaaaca cacgcetctag	gaaaacaatg	tccaacaggg cactgcaaag atctgccgtg	120
ctgtctgcgt ttgttctgt	tcgtgttct	ccggagacgg gaaagcaata	180
tgggataaaa aacagtacgt	gagtcggta	aacccaagtc agctgttcct ctatgacact	240
ttccctaaaa actttcctg	ggcggttggg	accggagcat ttcaagtgga agggagttgg	300
aagacagatg gaagaggacc	ctcgatctgg	gatcggtacg tctactcaca cctgagaggt	360
gtcaacggca cagacagatc	cactgacagt	tacatcttc tggaaaaaga cttgttgct	420
ctggatttt taggagttt	tttttatcag	ttctcaatct cctggccacg gttgttccc	480
aatggaaacag tagcagcagt	gaatgcgaa	ggtctccggt actaccgtgc acttctggac	540
tcgctggta ttaggaatat	cgagccatt	gttaccttgc accattggg tttgcctctg	600
acgctccagg aagaatatgg	gggctggaaa	aatgcaacta tgatagatct cttcaacgac	660
tatgccacat actgcttcca	gacccttgg	gaccgtgtca aatattggat tacaattcac	720
aacccttacc ttgttgctt	gcatggg	ggcacaggtt tgcatgcacc aggagagaag	780
ggaaatttaa cagctgtcta	cactgtggg	cacaacctgt tcaaggcaca ttcaaggatg	840
tggcataact acgacaaaaaa	cttccgcct	catcagaagg gttggcttc catcaccttgc	900
gggtcccatt gnatagagcc	aaacagaaca	gacaacatgg aggacgtgt caactgcccag	960
cactccatgt cctctgtgt	tggatggttc	gccaacccca tccacgggca cggcgactac	1020
cctgagttca tgaagacggg	cgccatgtac	cccgagtttct ctgaggcaga gaaggaggag	1080
gtgagggca cggctgattt	cttgcctt	tccttcggc ccaacaactt caggccctca	1140
aacaccgtgg tgaaaatggg	acaaaatgt	tcaactcaact taaggcaggt gctgaactgg	1200
attaaactgg aatacgtat	ccctcaaatt	ttgatttcggtt agaacggctg gttcacagat	1260
agctatataa agacagagga	caccacggcc	atctacatgt tgaagaattt cctaaaccag	1320
gttcttaag caataaaaatt	tgtgaaatc	cgcggttttgc gttatacggc ctggactctc	1380
ctggatggct ttgagtgca	ggatgcctat	acgacccgac gagggctgtt ttatgtggac	1440
ttaaacatgt agcagaaaaga	gaggaaaccc	aagtccctcggtt ctcattacta caagcagatc	1500
atacaagaca acggcccccc	tttggaaagag	tccacgcccag acatgaaggg tcggttcccc	1560
tgtgattttt ctgggggagt	cactgagtct	gttcttaagc ccgagtttac ggtctccctcc	1620
ccgcagtttta cccatctca	cctgtatgt	tggaatgtca ctggcaacag attgtctac	1680
cgagtggaaag gggtaaggct	gaaaacaaga	ccatcccagt gcacagatata tggatggatc	1740
aaaaaacggat tgaaaatgtt	ggcaaaaatg	aaagtccacc actaccagg tgcgttggac	1800
tggacctcta tcctccac	tggcaatctg	tccaaagtta acagacaatgtt gttaaggat	1860
tatagggtgt tggtgagcga	aggactgaag	ctgggcgtct tccccatggt gacgttgtac	1920
cacccaaccc actcccatct	cgccctcccc	ctgcccacttc tgagcgtgg ggggtggcta	1980
aacatgaaca cagccaaggc	cttccaggac	tacgctgacg tggcttccgg ggagttgggg	2040
gacttggta agctctggat	caccatcaat	gagcctaaca ggctgagtgatgtac	2100
cgcacgagta atgacacacta	ccgtgcagcc	cacaacctgt tgcgtccca tgcccaggtc	2160

US 9,475,856 B2

359**360**

-continued

```

tggcacctct atgataggca gtataggccg gtccagcatg gggctgtgtc gctgtcctta    2220
cattgcgact gggcagaacc tgccaacccc tttgtggatt cacactggaa ggcagccgag    2280
cgcttcctcc agtttgagat cgccctggtt gcagatccgc tcttcaagac tggcgactat    2340
ccatcggtta tgaaggaaa catcgccctcc aagaaccagg gagggctgtc tagctcagtc    2400
ctgccgcgct tcaccgcgaa ggagagcagg ctggtaagg gtaccgtcga cttctacgca    2460
ctgaaccact tcactacgag gtccgtgata cacaaggcgc tgaacaccaa ccgctcagtt    2520
geagacaggg acgtccagtt cctgcaggac atcaccgcgc taagctcgc cagccgcctg    2580
getgtaacac cctggggagt ggcgaagtc cttgcgtgaa tccggaggaa ctacagagac    2640
agggatatct acatcacagc caatggcattc gatgacctgg ctctagagga tgatcagatc    2700
cgaaagtact acttggagaa gtatgtccag gaggctctga aagcatatct cattgacaag    2760
gtcaaaatca aaggctacta tgcatcaaa ctgactgaag agaaatctaa gcctagattt    2820
ggattttca cctctgactt cagagctaag tccatctgtcc agtttacag caagctgatc    2880
agcagcagtg gcctccccgc tgagaacaga agtccctgcgt gtggcagcc tgccgaagac    2940
acagactgca ccatttgcgc atttctcggt gagaagaaac cactcatctt cttcggttgc    3000
tgcttcatct ccactctggc tggactgcta tccatcacccg tttttcatca tcaaaagaga    3060
agaaaattcc agaaagcaag gaacctacaa aatataccat tgaagaaagg ccacagcaga    3120
gttttcagct aa                                         3132

```

<210> SEQ ID NO 221

<211> LENGTH: 822

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 221

Met	Trp	Ser	Trp	Lys	Cys	Leu	Leu	Phe	Trp	Ala	Val	Leu	Val	Thr	Ala
1						5			10			15			

Thr	Leu	Cys	Thr	Ala	Arg	Pro	Ser	Pro	Thr	Leu	Pro	Glu	Gln	Ala	Gln
			20				25			30					

Pro	Trp	Gly	Ala	Pro	Val	Glu	Val	Glu	Ser	Phe	Leu	Val	His	Pro	Gly
	35				40			45							

Asp	Leu	Leu	Gln	Leu	Arg	Cys	Arg	Leu	Arg	Asp	Asp	Val	Gln	Ser	Ile
	50				55			60							

Asn	Trp	Leu	Arg	Asp	Gly	Val	Gln	Leu	Ala	Glu	Ser	Asn	Arg	Thr	Arg
65					70			75		80					

Ile	Thr	Gly	Glu	Glu	Val	Glu	Val	Gln	Asp	Ser	Val	Pro	Ala	Asp	Ser
	85					90		95							

Gly	Leu	Tyr	Ala	Cys	Val	Thr	Ser	Ser	Pro	Ser	Gly	Ser	Asp	Thr	Thr
	100				105			110							

Tyr	Phe	Ser	Val	Asn	Val	Ser	Asp	Ala	Leu	Pro	Ser	Ser	Glu	Asp	Asp
	115				120			125							

Asp	Asp	Asp	Asp	Ser	Ser	Ser	Glu	Glu	Lys	Asp	Asn	Thr			
	130				135			140							

Lys	Pro	Asn	Arg	Met	Pro	Val	Ala	Pro	Tyr	Trp	Thr	Ser	Pro	Glu	Lys
145					150			155		160					

Met	Glu	Lys	Lys	Leu	His	Ala	Val	Pro	Ala	Ala	Lys	Thr	Val	Lys	Phe
	165				170			175							

Lys	Cys	Pro	Ser	Ser	Gly	Thr	Pro	Asn	Pro	Thr	Leu	Arg	Trp	Leu	Lys
	180				185			190							

-continued

Asn Gly Lys Glu Phe Lys Pro Asp His Arg Ile Gly Gly Tyr Lys Val
 195 200 205
 Arg Tyr Ala Thr Trp Ser Ile Ile Met Asp Ser Val Val Pro Ser Asp
 210 215 220
 Lys Gly Asn Tyr Thr Cys Ile Val Glu Asn Glu Tyr Gly Ser Ile Asn
 225 230 235 240
 His Thr Tyr Gln Leu Asp Val Val Glu Arg Ser Pro His Arg Pro Ile
 245 250 255
 Leu Gln Ala Gly Leu Pro Ala Asn Lys Thr Val Ala Leu Gly Ser Asn
 260 265 270
 Val Glu Phe Met Cys Lys Val Tyr Ser Asp Pro Gln Pro His Ile Gln
 275 280 285
 Trp Leu Lys His Ile Glu Val Asn Gly Ser Lys Ile Gly Pro Asp Asn
 290 295 300
 Leu Pro Tyr Val Gln Ile Leu Lys Thr Ala Gly Val Asn Thr Thr Asp
 305 310 315 320
 Lys Glu Met Glu Val Leu His Leu Arg Asn Val Ser Phe Glu Asp Ala
 325 330 335
 Gly Glu Tyr Thr Cys Leu Ala Gly Asn Ser Ile Gly Leu Ser His His
 340 345 350
 Ser Ala Trp Leu Thr Val Leu Glu Ala Leu Glu Glu Arg Pro Ala Val
 355 360 365
 Met Thr Ser Pro Leu Tyr Leu Glu Ile Ile Tyr Cys Thr Gly Ala
 370 375 380
 Phe Leu Ile Ser Cys Met Val Gly Ser Val Ile Val Tyr Lys Met Lys
 385 390 395 400
 Ser Gly Thr Lys Lys Ser Asp Phe His Ser Gln Met Ala Val His Lys
 405 410 415
 Leu Ala Lys Ser Ile Pro Leu Arg Arg Gln Val Thr Val Ser Ala Asp
 420 425 430
 Ser Ser Ala Ser Met Asn Ser Gly Val Leu Leu Val Arg Pro Ser Arg
 435 440 445
 Leu Ser Ser Ser Gly Thr Pro Met Leu Ala Gly Val Ser Glu Tyr Glu
 450 455 460
 Leu Pro Glu Asp Pro Arg Trp Glu Leu Pro Arg Asp Arg Leu Val Leu
 465 470 475 480
 Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val Val Leu Ala Glu
 485 490 495
 Ala Ile Gly Leu Asp Lys Asp Lys Pro Asn Arg Val Thr Lys Val Ala
 500 505 510
 Val Lys Met Leu Lys Ser Asp Ala Thr Glu Lys Asp Leu Ser Asp Leu
 515 520 525
 Ile Ser Glu Met Glu Met Met Lys Met Ile Gly Lys His Lys Asn Ile
 530 535 540
 Ile Asn Leu Leu Gly Ala Cys Thr Gln Asp Gly Pro Leu Tyr Val Ile
 545 550 555 560
 Val Glu Tyr Ala Ser Lys Gly Asn Leu Arg Glu Tyr Leu Gln Ala Arg
 565 570 575
 Arg Pro Pro Gly Leu Glu Tyr Cys Tyr Asn Pro Ser His Asn Pro Glu
 580 585 590
 Glu Gln Leu Ser Ser Lys Asp Leu Val Ser Cys Ala Tyr Gln Val Ala
 595 600 605
 Arg Gly Met Glu Tyr Leu Ala Ser Lys Lys Cys Ile His Arg Asp Leu

US 9,475,856 B2

363**364**

-continued

610	615	620
Ala Ala Arg Asn Val Leu Val Thr Glu Asp Asn Val Met Lys Ile Ala		
625	630	635
Asp Phe Gly Leu Ala Arg Asp Ile His His Ile Asp Tyr Tyr Lys Lys		
645	650	655
Thr Thr Asn Gly Arg Leu Pro Val Lys Trp Met Ala Pro Glu Ala Leu		
660	665	670
Phe Asp Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser Phe Gly Val		
675	680	685
Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Ser Pro Tyr Pro Gly Val		
690	695	700
Pro Val Glu Glu Leu Phe Lys Leu Leu Lys Glu Gly His Arg Met Asp		
705	710	715
Lys Pro Ser Asn Cys Thr Asn Glu Leu Tyr Met Met Met Arg Asp Cys		
725	730	735
Trp His Ala Val Pro Ser Gln Arg Pro Thr Phe Lys Gln Leu Val Glu		
740	745	750
Asp Leu Asp Arg Ile Val Ala Leu Thr Ser Asn Gln Glu Tyr Leu Asp		
755	760	765
Leu Ser Met Pro Leu Asp Gln Tyr Ser Pro Ser Phe Pro Asp Thr Arg		
770	775	780
Ser Ser Thr Cys Ser Ser Gly Glu Asp Ser Val Phe Ser His Glu Pro		
785	790	795
Leu Pro Glu Pro Cys Leu Pro Arg His Pro Ala Gln Leu Ala Asn		
805	810	815
Gly Gly Leu Lys Arg Arg		
820		

<210> SEQ ID NO 222

<211> LENGTH: 2469

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 222

```

atgtggagct ggaagtgcct cctttctgg gctgtgtgg tcacagccac actctgcacc      60
gctaggccgt ccccaccc ttgcgttgcacaa gcccggccctt gggggcccccc tggaaatgt      120
gagtccttcc tggccaccc cggtgacccg ctgcagttc gctgtcggtt gcggggacat      180
gtgcagagca tcaactggct gccccgggggtt gtcgtggcc cggaaagccaa ccgcaccgc      240
atcacagggg aggagggttggaa ggtgcaggac tccgtgccc cagactccgg cctctatgt      300
tgcgttaaccca gcagccccctc gggcagtgtac accacctact tctccgtcaa tgttttagat      360
gtctccccc cctccggagga tggatgtat gatgtatgtt cctcttcaga ggaaaaagaa      420
acagataaca ccaaaacccaaa ccgttatgccc gtatgtccat attggacatc cccagaaaaag      480
atggaaaaga aattgtcatgc agtgcggctt gccaagacag tgaagttcaa atgccttcc      540
agtggggaccc caaaccacact gtcgtgtgg ttggaaaatg gcaaaagaatt caaacctgac      600
cacacatacc gaggttacaa ggtccgttat gccacctggaa gcatcataat ggactctgt      660
gtgccctctg acaaggccaa ctacacccgttccat gtcgtggata atgagttacgg cagcatcaac      720
cacacatacc agctggatgt cgtggagccgg tcccttcacc gggccatcct gcaaggccgg      780
ttggccggcca acaaaacactt gggccgtggatgtt gcaacgtgg agttcatgtt taagggttac      840
agtgacccgc agccgcacat ccagtggtca aagcacatcg aggtgaatgg gagcaagatt      900

```

US 9,475,856 B2

365

366

-continued

ggcccgagaca	acctgcctta	tgtccagatc	ttgaagactg	ctggaggtaa	taccaccgac	960
aaagagatgg	aggtgcttca	cttaagaaaat	gttccttttgc	aggacgcagg	ggagtatacgt	1020
tgcttggcgg	gtaaactctat	cggactctcc	catcaactctg	catgggttgc	cgttctggaa	1080
gcccttggaa	agaggccggc	agtgtatgacc	tcgcccctgt	acctggagat	catcatctat	1140
tgcacagggg	ctttcctcat	ctcctgcatg	gtggggtcgg	tatcgtcta	caagatgaag	1200
agtggatcca	agaagagtga	cttccacagc	cagatggctg	tgcacaagct	ggccaagagc	1260
atccctctgc	gcagacaggt	aacagtgtct	gctgactcca	gtgcatccat	gaactctggg	1320
gttcttctgg	ttcggccatc	acggctctcc	tccagtggga	ctcccatgct	agcaggggtc	1380
tctgagttat	agcttcccgaa	agaccctcgc	tggagactgc	ctcggtcacag	actggcttta	1440
ggcaaaacccc	tggagagaggg	ctgctttggg	caggtggtgt	tggcagaggg	tatcgggctg	1500
gacaaggaca	aacccaacccg	tgtgacccaaat	gtggctgtga	agatgttga	gtcggtacga	1560
acagagaaag	acttgtcaga	cctgtatctca	gaaatggaga	tgtgaagat	gatcgggaaag	1620
cataagaata	tcatcaacact	gctgggggcc	tgcacgcagg	atggttccctt	gtatgtcatc	1680
gtggagttat	cctccaaggg	caacctgcgg	gagtacctgc	aggccccggag	gccccccagg	1740
ctggaataact	gtcataaaccc	cagccacaac	ccagaggagc	agctctccctc	caaggacac	1800
gtgtcctgcg	cctaccagg	ggccccggc	atggagttatc	tggctccaa	gaagtgacata	1860
caccgagacc	tggcagccag	aatgtccctg	gtgacagagg	acaatgtat	gaagatagca	1920
gactttggcc	tgcacacgg	cattcaccac	atcgactact	ataaaaaagac	aaccaacccgc	1980
cgactgcctg	tgaagtggat	ggcacccggag	gcattatttg	accggatcta	caccacccag	2040
agtgtatgt	ggtctttcgg	ggtgctctgc	tggagatct	tcactctggg	cggctccccca	2100
taccccggtt	tgcctgtgga	ggaacttttc	aagctgtga	aggagggtca	ccgcatacgac	2160
aagcccgat	actgcaccaa	cgagctgtac	atgtatgtgc	gggactgtgc	gcatgcagtg	2220
ccctcacaga	gaccacccctt	caagcagctg	gtggaaagacc	tggaccgcac	cgtggcccttgc	2280
acctccaaacc	aggagttact	ggacactgtcc	atgccccctgg	accagttactc	ccccagcttt	2340
cccgacaccc	ggagctctac	gtgctctca	ggggaggatt	ccgttcttc	tcatgagccg	2400
ctgccccgagg	agccctgcct	gccccgacac	ccagcccgac	ttgccaatgg	cggaactcaaa	2460
cgccgctga						2469

<210> SEQ ID NO 223

<211> LENGTH: 251

<212> TYPE: PRT

<213> ORGANTSM: Homo sapiens

<400> SEQUENCE: 223

Met Leu Gly Ala Arg Leu Arg Leu Trp Val Cys Ala Leu Cys Ser Val
1 5 10 15

Cys Ser Met Ser Val Leu Arg Ala Tyr Pro Asn Ala Ser Pro Leu Leu
20 25 30

Gly Ser Ser Trp Gly Gly Leu Ile His Leu Tyr Thr Ala Thr Ala Arg
 35 40 45

Pro	His	Gln	Thr	Ile	Tyr	Ser	Ala	Leu	Met	Ile	Arg	Ser	Glu	Asp	Ala
65				70						75					80

Gly Phe Val Val Ile Thr Gly Val Met Ser Arg Arg Tyr Leu Cys Met
85 90 95

-continued

Asp Phe Arg Gly Asn Ile Phe Gly Ser His Tyr Phe Asp Pro Glu Asn
 100 105 110
 Cys Arg Phe Gln His Gln Thr Leu Glu Asn Gly Tyr Asp Val Tyr His
 115 120 125
 Ser Pro Gln Tyr His Phe Leu Val Ser Leu Gly Arg Ala Lys Arg Ala
 130 135 140
 Phe Leu Pro Gly Met Asn Pro Pro Pro Tyr Ser Gln Phe Leu Ser Arg
 145 150 155 160
 Arg Asn Glu Ile Pro Leu Ile His Phe Asn Thr Pro Ile Pro Arg Arg
 165 170 175
 His Thr Arg Ser Ala Glu Asp Asp Ser Glu Arg Asp Pro Leu Asn Val
 180 185 190
 Leu Lys Pro Arg Ala Arg Met Thr Pro Ala Pro Ala Ser Cys Ser Gln
 195 200 205
 Glu Leu Pro Ser Ala Glu Asp Asn Ser Pro Met Ala Ser Asp Pro Leu
 210 215 220
 Gly Val Val Arg Gly Gly Arg Val Asn Thr His Ala Gly Gly Thr Gly
 225 230 235 240
 Pro Glu Gly Cys Arg Pro Phe Ala Lys Phe Ile
 245 250

<210> SEQ ID NO 224

<211> LENGTH: 251

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 224

Met Leu Gly Ala Arg Leu Arg Leu Trp Val Cys Ala Leu Cys Ser Val
 1 5 10 15
 Cys Ser Met Ser Val Leu Arg Ala Tyr Pro Asn Ala Ser Pro Leu Leu
 20 25 30
 Gly Ser Ser Trp Gly Gly Leu Ile His Leu Tyr Thr Ala Thr Ala Arg
 35 40 45
 Asn Ser Tyr His Leu Gln Ile His Lys Asn Gly His Val Asp Gly Ala
 50 55 60
 Pro His Gln Thr Ile Tyr Ser Ala Leu Met Ile Arg Ser Glu Asp Ala
 65 70 75 80
 Gly Phe Val Val Ile Thr Gly Val Met Ser Arg Arg Tyr Leu Cys Thr
 85 90 95
 Asp Phe Arg Gly Asn Ile Phe Gly Ser His Tyr Phe Asp Pro Glu Asn
 100 105 110
 Cys Arg Phe Gln His Gln Thr Leu Glu Asn Gly Tyr Asp Val Tyr His
 115 120 125
 Ser Pro Gln Tyr His Phe Leu Val Ser Leu Gly Arg Ala Lys Arg Ala
 130 135 140
 Phe Leu Pro Gly Met Asn Pro Pro Pro Tyr Ser Gln Phe Leu Ser Arg
 145 150 155 160
 Arg Asn Glu Ile Pro Leu Ile His Phe Asn Thr Pro Ile Pro Arg Arg
 165 170 175
 His Thr Arg Ser Ala Glu Asp Asp Ser Glu Arg Asp Pro Leu Asn Val
 180 185 190
 Leu Lys Pro Arg Ala Arg Met Thr Pro Ala Pro Ala Ser Cys Ser Gln
 195 200 205
 Glu Leu Pro Ser Ala Glu Asp Asn Ser Pro Met Ala Ser Asp Pro Leu

US 9,475,856 B2

369

370

-continued

210 215 220

Gly Val Val Arg Gly Gly Arg Val Asn Thr His Ala Gly Gly Thr Gly
 225 230 235 240

Pro Glu Gly Cys Arg Pro Phe Ala Lys Phe Ile
 245 250

<210> SEQ ID NO 225

<211> LENGTH: 825

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 225

Met Trp Ser Trp Lys Cys Leu Leu Phe Trp Ala Val Leu Val Thr Ala
 1 5 10 15

Thr Leu Cys Thr Ala Arg Pro Ser Pro Thr Leu Pro Glu Gln Ala Gln
 20 25 30

Pro Trp Gly Ala Pro Val Glu Val Glu Ser Phe Leu Val His Pro Gly
 35 40 45

Asp Leu Leu Gln Leu Arg Cys Arg Leu Arg Asp Asp Val Gln Ser Ile
 50 55 60

Asn Trp Leu Arg Asp Gly Val Gln Leu Ala Glu Ser Asn Arg Thr Arg
 65 70 75 80

Ile Thr Gly Glu Glu Val Glu Val Gln Asp Ser Val Pro Ala Asp Ser
 85 90 95

Gly Leu Tyr Ala Cys Val Thr Ser Ser Pro Ser Gly Ser Asp Thr Thr
 100 105 110

Tyr Phe Ser Val Asn Val Ser Val Pro Ile Asp Ala Leu Pro Ser Ser
 115 120 125

Glu Asp Asp Asp Asp Asp Asp Ser Ser Ser Glu Glu Lys Glu Thr
 130 135 140

Asp Asn Thr Lys Pro Asn Pro Val Ala Pro Tyr Trp Thr Ser Pro Glu
 145 150 155 160

Lys Met Glu Lys Lys Leu His Ala Val Pro Ala Ala Lys Thr Val Lys
 165 170 175

Phe Lys Cys Pro Ser Ser Gly Thr Pro Asn Pro Thr Leu Arg Trp Leu
 180 185 190

Lys Asn Gly Lys Glu Phe Lys Pro Asp His Arg Ile Gly Tyr Lys
 195 200 205

Val Arg Tyr Ala Thr Trp Ser Ile Ile Met Asp Ser Val Val Pro Ser
 210 215 220

Asp Lys Gly Asn Tyr Thr Cys Ile Val Glu Asn Glu Tyr Gly Ser Ile
 225 230 235 240

Asn His Thr Tyr Gln Leu Asp Val Val Glu Arg Ser Pro His Arg Pro
 245 250 255

Ile Leu Gln Ala Gly Leu Pro Ala Asn Lys Thr Val Ala Leu Gly Ser
 260 265 270

Asn Val Glu Phe Met Cys Lys Val Tyr Ser Asp Pro Gln Pro His Ile
 275 280 285

Gln Trp Leu Lys His Ile Glu Val Asn Gly Ser Lys Ile Gly Pro Asp
 290 295 300

Asn Leu Pro Tyr Val Gln Ile Leu Lys His Ser Gly Ile Asn Ser Ser
 305 310 315 320

Asp Ala Glu Val Leu Thr Leu Phe Asn Val Thr Glu Ala Gln Ser Gly
 325 330 335

-continued

Glu Tyr Val Cys Lys Val Ser Asn Tyr Ile Gly Glu Ala Asn Gln Ser
340 345 350

Ala Trp Leu Thr Val Thr Arg Pro Val Ala Lys Ala Leu Glu Glu Arg
355 360 365

Pro Ala Val Met Thr Ser Pro Leu Tyr Leu Glu Ile Ile Ile Tyr Cys
370 375 380

Thr Gly Ala Phe Leu Ile Ser Cys Met Val Gly Ser Val Ile Val Tyr
385 390 395 400

Lys Met Lys Ser Gly Thr Lys Ser Asp Phe His Ser Gln Met Ala
405 410 415

Val His Lys Leu Ala Lys Ser Ile Pro Leu Arg Arg Gln Val Thr Val
420 425 430

Ser Ala Asp Ser Ser Ala Ser Met Asn Ser Gly Val Leu Leu Val Arg
435 440 445

Pro Ser Arg Leu Ser Ser Gly Thr Pro Met Leu Ala Gly Val Ser
450 455 460

Glu Tyr Glu Leu Pro Glu Asp Pro Arg Trp Glu Leu Pro Arg Asp Arg
465 470 475 480

Leu Val Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val Val
485 490 495

Leu Ala Glu Ala Ile Gly Leu Asp Lys Asp Lys Pro Asn Arg Val Thr
500 505 510

Lys Val Ala Val Lys Met Leu Lys Ser Asp Ala Thr Glu Lys Asp Leu
515 520 525

Ser Asp Leu Ile Ser Glu Met Glu Met Met Lys Met Ile Gly Lys His
530 535 540

Lys Asn Ile Ile Asn Leu Leu Gly Ala Cys Thr Gln Asp Gly Pro Leu
545 550 555 560

Tyr Val Ile Val Glu Tyr Ala Ser Lys Gly Asn Leu Arg Glu Tyr Leu
565 570 575

Gln Ala Arg Arg Pro Pro Gly Leu Glu Tyr Cys Tyr Asn Pro Ser His
580 585 590

Asn Pro Glu Glu Gln Leu Ser Ser Lys Asp Leu Val Ser Cys Ala Tyr
595 600 605

Gln Val Ala Arg Gly Met Glu Tyr Leu Ala Ser Lys Lys Cys Ile His
610 615 620

Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asp Asn Val Met
625 630 635 640

Lys Ile Ala Asp Phe Gly Leu Ala Arg Asp Ile His His Ile Asp Tyr
645 650 655

Tyr Lys Lys Thr Thr Asn Gly Arg Leu Pro Val Lys Trp Met Ala Pro
660 665 670

Glu Ala Leu Phe Asp Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser
675 680 685

Phe Gly Val Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Ser Pro Tyr
690 695 700

Pro Gly Val Pro Val Glu Glu Leu Phe Lys Leu Leu Lys Glu Gly His
705 710 715 720

Arg Met Asp Lys Pro Ser Asn Cys Thr Asn Glu Leu Tyr Met Met Met
725 730 735

Arg Asp Cys Trp His Ala Val Pro Ser Gln Arg Pro Thr Phe Lys Gln
740 745 750

Leu Val Glu Asp Leu Asp Arg Ile Val Ala Leu Thr Ser Asn Gln Glu

US 9,475,856 B2

373

374

-continued

755

760

765

Tyr Leu Asp Leu Ser Met Pro Leu Asp Gln Tyr Ser Pro Ser Phe Pro
 770 775 780

Asp Thr Arg Ser Ser Thr Cys Ser Ser Gly Glu Asp Ser Val Phe Ser
 785 790 795 800

His Glu Pro Leu Pro Glu Glu Pro Cys Leu Pro Arg His Pro Ala Gln
 805 810 815

Leu Ala Asn Gly Gly Leu Lys Arg Arg
 820 825

<210> SEQ ID NO 226

<211> LENGTH: 821

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 226

Met Val Ser Trp Gly Arg Phe Ile Cys Leu Val Val Val Thr Met Ala
 1 5 10 15

Thr Leu Ser Leu Ala Arg Pro Ser Phe Ser Leu Val Glu Asp Thr Thr
 20 25 30

Leu Glu Pro Glu Glu Pro Pro Thr Lys Tyr Gln Ile Ser Gln Pro Glu
 35 40 45

Val Tyr Val Ala Ala Pro Gly Glu Ser Leu Glu Val Arg Cys Leu Leu
 50 55 60

Lys Asp Ala Ala Val Ile Ser Trp Thr Lys Asp Gly Val His Leu Gly
 65 70 75 80

Pro Asn Asn Arg Thr Val Leu Ile Gly Glu Tyr Leu Gln Ile Lys Gly
 85 90 95

Ala Thr Pro Arg Asp Ser Gly Leu Tyr Ala Cys Thr Ala Ser Arg Thr
 100 105 110

Val Asp Ser Glu Thr Trp Tyr Phe Met Val Asn Val Thr Asp Ala Ile
 115 120 125

Ser Ser Gly Asp Asp Glu Asp Asp Thr Asp Gly Ala Glu Asp Phe Val
 130 135 140

Ser Glu Asn Ser Asn Asn Lys Arg Ala Pro Tyr Trp Thr Asn Thr Glu
 145 150 155 160

Lys Met Glu Lys Arg Leu His Ala Val Pro Ala Ala Asn Thr Val Lys
 165 170 175

Phe Arg Cys Pro Ala Gly Gly Asn Pro Met Pro Thr Met Arg Trp Leu
 180 185 190

Lys Asn Gly Lys Glu Phe Lys Gln Glu His Arg Ile Gly Tyr Lys
 195 200 205

Val Arg Asn Gln His Trp Ser Leu Ile Met Glu Ser Val Val Pro Ser
 210 215 220

Asp Lys Gly Asn Tyr Thr Cys Val Val Glu Asn Glu Tyr Gly Ser Ile
 225 230 235 240

Asn His Thr Tyr His Leu Asp Val Val Glu Arg Ser Pro His Arg Pro
 245 250 255

Ile Leu Gln Ala Gly Leu Pro Ala Asn Ala Ser Thr Val Val Gly Gly
 260 265 270

Asp Val Glu Phe Val Cys Lys Val Tyr Ser Asp Ala Gln Pro His Ile
 275 280 285

Gln Trp Ile Lys His Val Glu Lys Asn Gly Ser Lys Tyr Gly Pro Asp
 290 295 300

-continued

Gly Leu Pro Tyr Leu Lys Val Leu Lys Ala Ala Gly Val Asn Thr Thr
305 310 315 320

Asp Lys Glu Ile Glu Val Leu Tyr Ile Arg Asn Val Thr Phe Glu Asp
325 330 335

Ala Gly Glu Tyr Thr Cys Leu Ala Gly Asn Ser Ile Gly Ile Ser Phe
340 345 350

His Ser Ala Trp Leu Thr Val Leu Pro Ala Pro Gly Arg Glu Lys Glu
355 360 365

Ile Thr Ala Ser Pro Asp Tyr Leu Glu Ile Ala Ile Tyr Cys Ile Gly
370 375 380

Val Phe Leu Ile Ala Cys Met Val Val Thr Val Ile Leu Cys Arg Met
385 390 395 400

Lys Asn Thr Thr Lys Lys Pro Asp Phe Ser Ser Gln Pro Ala Val His
405 410 415

Lys Leu Thr Lys Arg Ile Pro Leu Arg Arg Gln Val Thr Val Ser Ala
420 425 430

Glu Ser Ser Ser Met Asn Ser Asn Thr Pro Leu Val Arg Ile Thr
435 440 445

Thr Arg Leu Ser Ser Thr Ala Asp Thr Pro Met Leu Ala Gly Val Ser
450 455 460

Glu Tyr Glu Leu Pro Glu Asp Pro Lys Trp Glu Phe Pro Arg Asp Lys
465 470 475 480

Leu Thr Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val Val
485 490 495

Met Ala Glu Ala Val Gly Ile Asp Lys Asp Lys Pro Lys Glu Ala Val
500 505 510

Thr Val Ala Val Lys Met Leu Lys Asp Asp Ala Thr Glu Lys Asp Leu
515 520 525

Ser Asp Leu Val Ser Glu Met Glu Met Met Lys Met Ile Gly Lys His
530 535 540

Lys Asn Ile Ile Asn Leu Leu Gly Ala Cys Thr Gln Asp Gly Pro Leu
545 550 555 560

Tyr Val Ile Val Glu Tyr Ala Ser Lys Gly Asn Leu Arg Glu Tyr Leu
565 570 575

Arg Ala Arg Arg Pro Pro Gly Met Glu Tyr Ser Tyr Asp Ile Asn Arg
580 585 590

Val Pro Glu Glu Gln Met Thr Phe Lys Asp Leu Val Ser Cys Thr Tyr
595 600 605

Gln Leu Ala Arg Gly Met Glu Tyr Leu Ala Ser Gln Lys Cys Ile His
610 615 620

Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asn Asn Val Met
625 630 635 640

Lys Ile Ala Asp Phe Gly Leu Ala Arg Asp Ile Asn Asn Ile Asp Tyr
645 650 655

Tyr Lys Lys Thr Thr Asn Gly Arg Leu Pro Val Lys Trp Met Ala Pro
660 665 670

Glu Ala Leu Phe Asp Arg Val Tyr Thr His Gln Ser Asp Val Trp Ser
675 680 685

Phe Gly Val Leu Met Trp Glu Ile Phe Thr Leu Gly Gly Ser Pro Tyr
690 695 700

Pro Gly Ile Pro Val Glu Glu Leu Phe Lys Leu Leu Lys Glu Gly His
705 710 715 720

Arg Met Asp Lys Pro Ala Asn Cys Thr Asn Glu Leu Tyr Met Met Met

US 9,475,856 B2

379**380**

-continued

Asp Val Glu Phe Val Cys Lys Val Tyr Ser Asp Ala Gln Pro His Ile
 275 280 285
 Gln Trp Ile Lys His Val Glu Lys Asn Gly Ser Lys Tyr Gly Pro Asp
 290 295 300
 Gly Leu Pro Tyr Leu Lys Val Leu Lys His Ser Gly Ile Asn Ser Ser
 305 310 315 320
 Asn Ala Glu Val Leu Ala Leu Phe Asn Val Thr Glu Ala Asp Ala Gly
 325 330 335
 Glu Tyr Ile Cys Lys Val Ser Asn Tyr Ile Gly Gln Ala Asn Gln Ser
 340 345 350
 Ala Trp Leu Thr Val Leu Pro Lys Gln Gln Ala Pro Gly Arg Glu Lys
 355 360 365
 Glu Ile Thr Ala Ser Pro Asp Tyr Leu Glu Ile Ala Ile Tyr Cys Ile
 370 375 380
 Gly Val Phe Leu Ile Ala Cys Met Val Val Thr Val Ile Leu Cys Arg
 385 390 395 400
 Met Lys Asn Thr Thr Lys Lys Pro Asp Phe Ser Ser Gln Pro Ala Val
 405 410 415
 His Lys Leu Thr Lys Arg Ile Pro Leu Arg Arg Gln Val Thr Val Ser
 420 425 430
 Ala Glu Ser Ser Ser Met Asn Ser Asn Thr Pro Leu Val Arg Ile
 435 440 445
 Thr Thr Arg Leu Ser Ser Thr Ala Asp Thr Pro Met Leu Ala Gly Val
 450 455 460
 Ser Glu Tyr Glu Leu Pro Glu Asp Pro Lys Trp Glu Phe Pro Arg Asp
 465 470 475 480
 Lys Leu Thr Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val
 485 490 495
 Val Met Ala Glu Ala Val Gly Ile Asp Lys Asp Lys Pro Lys Glu Ala
 500 505 510
 Val Thr Val Ala Val Lys Met Leu Lys Asp Asp Ala Thr Glu Lys Asp
 515 520 525
 Leu Ser Asp Leu Val Ser Glu Met Glu Met Met Lys Met Ile Gly Lys
 530 535 540
 His Lys Asn Ile Ile Asn Leu Leu Gly Ala Cys Thr Gln Asp Gly Pro
 545 550 555 560
 Leu Tyr Val Ile Val Glu Tyr Ala Ser Lys Gly Asn Leu Arg Glu Tyr
 565 570 575
 Leu Arg Ala Arg Arg Pro Pro Gly Met Glu Tyr Ser Tyr Asp Ile Asn
 580 585 590
 Arg Val Pro Glu Glu Gln Met Thr Phe Lys Asp Leu Val Ser Cys Thr
 595 600 605
 Tyr Gln Leu Ala Arg Arg Met Glu Tyr Leu Ala Ser Gln Lys Cys Ile
 610 615 620
 His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asn Asn Val
 625 630 635 640
 Met Lys Ile Ala Asp Phe Gly Leu Ala Arg Asp Ile Asn Asn Ile Asp
 645 650 655
 Tyr Tyr Lys Lys Thr Thr Asn Gly Arg Leu Pro Val Lys Trp Met Ala
 660 665 670
 Pro Glu Ala Leu Phe Asp Arg Val Tyr Thr His Gln Ser Asp Val Trp
 675 680 685
 Ser Phe Gly Val Leu Met Trp Glu Ile Phe Thr Leu Gly Ser Pro

US 9,475,856 B2

381**382**

-continued

690 695 700

Tyr Pro Gly Ile Pro Val Glu Glu Leu Phe Lys Leu Leu Lys Glu Gly
 705 710 715 720

His Arg Met Asp Lys Pro Ala Asn Cys Thr Asn Glu Leu Tyr Met Met
 725 730 735

Met Arg Asp Cys Trp His Ala Val Pro Ser Gln Arg Pro Thr Phe Lys
 740 745 750

Gln Leu Val Glu Asp Leu Asp Arg Ile Leu Thr Leu Thr Thr Asn Glu
 755 760 765

Glu Tyr Leu Asp Leu Ser Gln Pro Leu Glu Gln Tyr Ser Pro Ser Tyr
 770 775 780

Pro Asp Thr Arg Ser Ser Cys Ser Ser Gly Asp Asp Ser Val Phe Ser
 785 790 795 800

Pro Asp Pro Met Pro Tyr Glu Pro Cys Leu Pro Gln Tyr Pro His Ile
 805 810 815

Asn Gly Ser Val Lys Thr
 820

<210> SEQ ID NO 228

<211> LENGTH: 806

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 228

Met Gly Ala Pro Ala Cys Ala Leu Ala Leu Cys Val Ala Val Ala Ile
 1 5 10 15

Val Ala Gly Ala Ser Ser Glu Ser Leu Gly Thr Glu Gln Arg Val Val
 20 25 30

Gly Arg Ala Ala Glu Val Pro Gly Pro Glu Pro Gly Gln Gln Glu Gln
 35 40 45

Leu Val Phe Gly Ser Gly Asp Ala Val Glu Leu Ser Cys Pro Pro Pro
 50 55 60

Gly Gly Gly Pro Met Gly Pro Thr Val Trp Val Lys Asp Gly Thr Gly
 65 70 75 80

Leu Val Pro Ser Glu Arg Val Leu Val Gly Pro Gln Arg Leu Gln Val
 85 90 95

Leu Asn Ala Ser His Glu Asp Ser Gly Ala Tyr Ser Cys Arg Gln Arg
 100 105 110

Leu Thr Gln Arg Val Leu Cys His Phe Ser Val Arg Val Thr Asp Ala
 115 120 125

Pro Ser Ser Gly Asp Asp Glu Asp Gly Glu Asp Glu Ala Glu Asp Thr
 130 135 140

Gly Val Asp Thr Gly Ala Pro Tyr Trp Thr Arg Pro Glu Arg Met Asp
 145 150 155 160

Lys Lys Leu Leu Ala Val Pro Ala Ala Asn Thr Val Arg Phe Arg Cys
 165 170 175

Pro Ala Ala Gly Asn Pro Thr Pro Ser Ile Ser Trp Leu Lys Asn Gly
 180 185 190

Arg Glu Phe Arg Gly Glu His Arg Ile Gly Gly Ile Lys Leu Arg His
 195 200 205

Gln Gln Trp Ser Leu Val Met Glu Ser Val Val Pro Ser Asp Arg Gly
 210 215 220

Asn Tyr Thr Cys Val Val Glu Asn Lys Phe Gly Ser Ile Arg Gln Thr
 225 230 235 240

-continued

Tyr Thr Leu Asp Val Leu Glu Arg Ser Pro His Arg Pro Ile Leu Gln
245 250 255

Ala Gly Leu Pro Ala Asn Gln Thr Ala Val Leu Gly Ser Asp Val Glu
260 265 270

Phe His Cys Lys Val Tyr Ser Asp Ala Gln Pro His Ile Gln Trp Leu
275 280 285

Lys His Val Glu Val Asn Gly Ser Lys Val Gly Pro Asp Gly Thr Pro
290 295 300

Tyr Val Thr Val Leu Lys Thr Ala Gly Ala Asn Thr Thr Asp Lys Glu
305 310 315 320

Leu Glu Val Leu Ser Leu His Asn Val Thr Phe Glu Asp Ala Gly Glu
325 330 335

Tyr Thr Cys Leu Ala Gly Asn Ser Ile Gly Phe Ser His His Ser Ala
340 345 350

Trp Leu Val Val Leu Pro Ala Glu Glu Leu Val Glu Ala Asp Glu
355 360 365

Ala Gly Ser Val Tyr Ala Gly Ile Leu Ser Tyr Gly Val Gly Phe Phe
370 375 380

Leu Phe Ile Leu Val Val Ala Ala Val Thr Leu Cys Arg Leu Arg Ser
385 390 395 400

Pro Pro Lys Gly Leu Gly Ser Pro Thr Val His Lys Ile Ser Arg
405 410 415

Phe Pro Leu Lys Arg Gln Val Ser Leu Glu Ser Asn Ala Ser Met Ser
420 425 430

Ser Asn Thr Pro Leu Val Arg Ile Ala Arg Leu Ser Ser Gly Glu Gly
435 440 445

Pro Thr Leu Ala Asn Val Ser Glu Leu Glu Leu Pro Ala Asp Pro Lys
450 455 460

Trp Glu Leu Ser Arg Ala Arg Leu Thr Leu Gly Lys Pro Leu Gly Glu
465 470 475 480

Gly Cys Phe Gly Gln Val Val Met Ala Glu Ala Ile Gly Ile Asp Lys
485 490 495

Asp Arg Ala Ala Lys Pro Val Thr Val Ala Val Lys Met Leu Lys Asp
500 505 510

Asp Ala Thr Asp Lys Asp Leu Ser Asp Leu Val Ser Glu Met Glu Met
515 520 525

Met Lys Met Ile Gly Lys His Lys Asn Ile Ile Asn Leu Leu Gly Ala
530 535 540

Cys Thr Gln Gly Pro Leu Tyr Val Leu Val Glu Tyr Ala Ala Lys
545 550 555 560

Gly Asn Leu Arg Glu Phe Leu Arg Ala Arg Arg Pro Pro Gly Leu Asp
565 570 575

Tyr Ser Phe Asp Thr Cys Lys Pro Pro Glu Glu Gln Leu Thr Phe Lys
580 585 590

Asp Leu Val Ser Cys Ala Tyr Gln Val Ala Arg Gly Met Glu Tyr Leu
595 600 605

Ala Ser Gln Lys Cys Ile His Arg Asp Leu Ala Ala Arg Asn Val Leu
610 615 620

Val Thr Glu Asp Asn Val Met Lys Ile Ala Asp Phe Gly Leu Ala Arg
625 630 635 640

Asp Val His Asn Leu Asp Tyr Tyr Lys Lys Thr Thr Asn Gly Arg Leu
645 650 655

Pro Val Lys Trp Met Ala Pro Glu Ala Leu Phe Asp Arg Val Tyr Thr

US 9,475,856 B2

385**386**

-continued

660	665	670
His Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe		
675	680	685
Thr Leu Gly Gly Ser Pro Tyr Pro Gly Ile Pro Val Glu Glu Leu Phe		
690	695	700
Lys Leu Leu Lys Glu Gly His Arg Met Asp Lys Pro Ala Asn Cys Thr		
705	710	715
720		
His Asp Leu Tyr Met Ile Met Arg Glu Cys Trp His Ala Ala Pro Ser		
725	730	735
Gln Arg Pro Thr Phe Lys Gln Leu Val Glu Asp Leu Asp Arg Val Leu		
740	745	750
Thr Val Thr Ser Thr Asp Glu Tyr Leu Asp Leu Ser Ala Pro Phe Glu		
755	760	765
Gln Tyr Ser Pro Gly Gly Gln Asp Thr Pro Ser Ser Ser Ser Gly		
770	775	780
Asp Asp Ser Val Phe Ala His Asp Leu Leu Pro Pro Ala Pro Pro Ser		
785	790	795
800		
Ser Gly Gly Ser Arg Thr		
805		

<210> SEQ ID NO 229
<211> LENGTH: 808
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 229

Met Gly Ala Pro Ala Cys Ala Leu Ala Leu Cys Val Ala Val Ala Ile	10	15
1	5	
Val Ala Gly Ala Ser Ser Glu Ser Leu Gly Thr Glu Gln Arg Val Val		
20	25	30
Gly Arg Ala Ala Glu Val Pro Gly Pro Glu Pro Gly Gln Gln Glu Gln		
35	40	45
Leu Val Phe Gly Ser Gly Asp Ala Val Glu Leu Ser Cys Pro Pro Pro		
50	55	60
Gly Gly Gly Pro Met Gly Pro Thr Val Trp Val Lys Asp Gly Thr Gly		
65	70	75
80		
Leu Val Pro Ser Glu Arg Val Leu Val Gly Pro Gln Arg Leu Gln Val		
85	90	95
Leu Asn Ala Ser His Glu Asp Ser Gly Ala Tyr Ser Cys Arg Gln Arg		
100	105	110
Leu Thr Gln Arg Val Leu Cys His Phe Ser Val Arg Val Thr Asp Ala		
115	120	125
Pro Ser Ser Gly Asp Asp Glu Asp Gly Glu Asp Glu Ala Glu Asp Thr		
130	135	140
Gly Val Asp Thr Gly Ala Pro Tyr Trp Thr Arg Pro Glu Arg Met Asp		
145	150	155
160		
Lys Lys Leu Leu Ala Val Pro Ala Ala Asn Thr Val Arg Phe Arg Cys		
165	170	175
Pro Ala Ala Gly Asn Pro Thr Pro Ser Ile Ser Trp Leu Lys Asn Gly		
180	185	190
Arg Glu Phe Arg Gly Glu His Arg Ile Gly Gly Ile Lys Leu Arg His		
195	200	205
Gln Gln Trp Ser Leu Val Met Glu Ser Val Val Pro Ser Asp Arg Gly		
210	215	220

-continued

Asn Tyr Thr Cys Val Val Glu Asn Lys Phe Gly Ser Ile Arg Gln Thr
 225 230 235 240
 Tyr Thr Leu Asp Val Leu Glu Arg Ser Pro His Arg Pro Ile Leu Gln
 245 250 255
 Ala Gly Leu Pro Ala Asn Gln Thr Ala Val Leu Gly Ser Asp Val Glu
 260 265 270
 Phe His Cys Lys Val Tyr Ser Asp Ala Gln Pro His Ile Gln Trp Leu
 275 280 285
 Lys His Val Glu Val Asn Gly Ser Lys Val Gly Pro Asp Gly Thr Pro
 290 295 300
 Tyr Val Thr Val Leu Lys Ser Trp Ile Ser Glu Ser Val Glu Ala Asp
 305 310 315 320
 Val Arg Leu Arg Leu Ala Asn Val Ser Glu Arg Asp Gly Gly Glu Tyr
 325 330 335
 Leu Cys Arg Ala Thr Asn Phe Ile Gly Val Ala Glu Lys Ala Phe Trp
 340 345 350
 Leu Ser Val His Gly Pro Arg Ala Ala Glu Glu Leu Val Glu Ala
 355 360 365
 Asp Glu Ala Gly Ser Val Tyr Ala Gly Ile Leu Ser Tyr Gly Val Gly
 370 375 380
 Phe Phe Leu Phe Ile Leu Val Val Ala Ala Val Thr Leu Cys Arg Leu
 385 390 395 400
 Arg Ser Pro Pro Lys Lys Gly Leu Gly Ser Pro Thr Val His Lys Ile
 405 410 415
 Ser Arg Phe Pro Leu Lys Arg Gln Val Ser Leu Glu Ser Asn Ala Ser
 420 425 430
 Met Ser Ser Asn Thr Pro Leu Val Arg Ile Ala Arg Leu Ser Ser Gly
 435 440 445
 Glu Gly Pro Thr Leu Ala Asn Val Ser Glu Leu Glu Leu Pro Ala Asp
 450 455 460
 Pro Lys Trp Glu Leu Ser Arg Ala Arg Leu Thr Leu Gly Lys Pro Leu
 465 470 475 480
 Gly Glu Gly Cys Phe Gly Gln Val Val Met Ala Glu Ala Ile Gly Ile
 485 490 495
 Asp Lys Asp Arg Ala Ala Lys Pro Val Thr Val Ala Val Lys Met Leu
 500 505 510
 Lys Asp Asp Ala Thr Asp Lys Asp Leu Ser Asp Leu Val Ser Glu Met
 515 520 525
 Glu Met Met Lys Met Ile Gly Lys His Lys Asn Ile Ile Asn Leu Leu
 530 535 540
 Gly Ala Cys Thr Gln Gly Pro Leu Tyr Val Leu Val Glu Tyr Ala
 545 550 555 560
 Ala Lys Gly Asn Leu Arg Glu Phe Leu Arg Ala Arg Arg Pro Pro Gly
 565 570 575
 Leu Asp Tyr Ser Phe Asp Thr Cys Lys Pro Pro Glu Glu Gln Leu Thr
 580 585 590
 Phe Lys Asp Leu Val Ser Cys Ala Tyr Gln Val Ala Arg Gly Met Glu
 595 600 605
 Tyr Leu Ala Ser Gln Lys Cys Ile His Arg Asp Leu Ala Ala Arg Asn
 610 615 620
 Val Leu Val Thr Glu Asp Asn Val Met Lys Ile Ala Asp Phe Gly Leu
 625 630 635 640
 Ala Arg Asp Val His Asn Leu Asp Tyr Tyr Lys Lys Thr Thr Asn Gly

US 9,475,856 B2

389**390**

-continued

645	650	655
Arg Leu Pro Val Lys Trp Met Ala Pro Glu Ala Leu Phe Asp Arg Val		
660	665	670
Tyr Thr His Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu		
675	680	685
Ile Phe Thr Leu Gly Gly Ser Pro Tyr Pro Gly Ile Pro Val Glu Glu		
690	695	700
Leu Phe Lys Leu Leu Lys Glu Gly His Arg Met Asp Lys Pro Ala Asn		
705	710	715
Cys Thr His Asp Leu Tyr Met Ile Met Arg Glu Cys Trp His Ala Ala		
725	730	735
Pro Ser Gln Arg Pro Thr Phe Lys Gln Leu Val Glu Asp Leu Asp Arg		
740	745	750
Val Leu Thr Val Thr Ser Thr Asp Glu Tyr Leu Asp Leu Ser Ala Pro		
755	760	765
Phe Glu Gln Tyr Ser Pro Gly Gly Gln Asp Thr Pro Ser Ser Ser Ser		
770	775	780
Ser Gly Asp Asp Ser Val Phe Ala His Asp Leu Leu Pro Pro Ala Pro		
785	790	795
Pro Ser Ser Gly Gly Ser Arg Thr		
805		

<210> SEQ ID NO 230

<211> LENGTH: 802

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 230

Met Arg Leu Leu Leu Ala Leu Leu Gly Val Leu Leu Ser Val Pro Gly		
1	5	10
15		
Pro Pro Val Leu Ser Leu Glu Ala Ser Glu Glu Val Glu Leu Glu Pro		
20	25	30
30		
Cys Leu Ala Pro Ser Leu Glu Gln Gln Glu Gln Glu Leu Thr Val Ala		
35	40	45
45		
Leu Gly Gln Pro Val Arg Leu Cys Cys Gly Arg Ala Glu Arg Gly Gly		
50	55	60
60		
His Trp Tyr Lys Glu Gly Ser Arg Leu Ala Pro Ala Gly Arg Val Arg		
65	70	75
75		
80		
Gly Trp Arg Gly Arg Leu Glu Ile Ala Ser Phe Leu Pro Glu Asp Ala		
85	90	95
95		
Gly Arg Tyr Leu Cys Leu Ala Arg Gly Ser Met Ile Val Leu Gln Asn		
100	105	110
110		
Leu Thr Leu Ile Thr Gly Asp Ser Leu Thr Ser Ser Asn Asp Asp Glu		
115	120	125
125		
Asp Pro Lys Ser His Arg Asp Pro Ser Asn Arg His Ser Tyr Pro Gln		
130	135	140
140		
Gln Ala Pro Tyr Trp Thr His Pro Gln Arg Met Glu Lys Lys Leu His		
145	150	155
155		
160		
Ala Val Pro Ala Gly Asn Thr Val Lys Phe Arg Cys Pro Ala Ala Gly		
165	170	175
175		
Asn Pro Thr Pro Thr Ile Arg Trp Leu Lys Asp Gly Gln Ala Phe His		
180	185	190
190		
Gly Glu Asn Arg Ile Gly Gly Ile Arg Leu Arg His Gln His Trp Ser		
195	200	205
205		

US 9,475,856 B2

391**392**

-continued

Leu Val Met Glu Ser Val Val Pro Ser Asp Arg Gly Thr Tyr Thr Cys
 210 215 220
 Leu Val Glu Asn Ala Val Gly Ser Ile Arg Tyr Asn Tyr Leu Leu Asp
 225 230 235 240
 Val Leu Glu Arg Ser Pro His Arg Pro Ile Leu Gln Ala Gly Leu Pro
 245 250 255
 Ala Asn Thr Thr Ala Val Val Gly Ser Asp Val Glu Leu Leu Cys Lys
 260 265 270
 Val Tyr Ser Asp Ala Gln Pro His Ile Gln Trp Leu Lys His Ile Val
 275 280 285
 Ile Asn Gly Ser Ser Phe Gly Ala Asp Gly Phe Pro Tyr Val Gln Val
 290 295 300
 Leu Lys Thr Ala Asp Ile Asn Ser Ser Glu Val Glu Val Leu Tyr Leu
 305 310 315 320
 Arg Asn Val Ser Ala Glu Asp Ala Gly Glu Tyr Thr Cys Leu Ala Gly
 325 330 335
 Asn Ser Ile Gly Leu Ser Tyr Gln Ser Ala Trp Leu Thr Val Leu Pro
 340 345 350
 Glu Glu Asp Pro Thr Trp Thr Ala Ala Ala Pro Glu Ala Arg Tyr Thr
 355 360 365
 Asp Ile Ile Leu Tyr Ala Ser Gly Ser Leu Ala Leu Ala Val Leu Leu
 370 375 380
 Leu Leu Ala Gly Leu Tyr Arg Gly Gln Ala Leu His Gly Arg His Pro
 385 390 395 400
 Arg Pro Pro Ala Thr Val Gln Lys Leu Ser Arg Phe Pro Leu Ala Arg
 405 410 415
 Gln Phe Ser Leu Glu Ser Gly Ser Ser Gly Lys Ser Ser Ser Leu
 420 425 430
 Val Arg Gly Val Arg Leu Ser Ser Ser Gly Pro Ala Leu Leu Ala Gly
 435 440 445
 Leu Val Ser Leu Asp Leu Pro Leu Asp Pro Leu Trp Glu Phe Pro Arg
 450 455 460
 Asp Arg Leu Val Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln
 465 470 475 480
 Val Val Arg Ala Glu Ala Phe Gly Met Asp Pro Ala Arg Pro Asp Gln
 485 490 495
 Ala Ser Thr Val Ala Val Lys Met Leu Lys Asp Asn Ala Ser Asp Lys
 500 505 510
 Asp Leu Ala Asp Leu Val Ser Glu Met Glu Val Met Lys Leu Ile Gly
 515 520 525
 Arg His Lys Asn Ile Ile Asn Leu Leu Gly Val Cys Thr Gln Glu Gly
 530 535 540
 Pro Leu Tyr Val Ile Val Glu Cys Ala Ala Lys Gly Asn Leu Arg Glu
 545 550 555 560
 Phe Leu Arg Ala Arg Arg Pro Pro Gly Pro Asp Leu Ser Pro Asp Gly
 565 570 575
 Pro Arg Ser Ser Glu Gly Pro Leu Ser Phe Pro Val Leu Val Ser Cys
 580 585 590
 Ala Tyr Gln Val Ala Arg Gly Met Gln Tyr Leu Glu Ser Arg Lys Cys
 595 600 605
 Ile His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asp Asn
 610 615 620
 Val Met Lys Ile Ala Asp Phe Gly Leu Ala Arg Gly Val His His Ile

US 9,475,856 B2

393**394**

-continued

625	630	635	640
Asp Tyr Tyr Lys Lys Thr Ser Asn Gly Arg Leu Pro Val Lys Trp Met			
645	650	655	
Ala Pro Glu Ala Leu Phe Asp Arg Val Tyr Thr His Gln Ser Asp Val			
660	665	670	
Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Ser			
675	680	685	
Pro Tyr Pro Gly Ile Pro Val Glu Glu Leu Phe Ser Leu Leu Arg Glu			
690	695	700	
Gly His Arg Met Asp Arg Pro Pro His Cys Pro Pro Glu Leu Tyr Gly			
705	710	715	720
Leu Met Arg Glu Cys Trp His Ala Ala Pro Ser Gln Arg Pro Thr Phe			
725	730	735	
Lys Gln Leu Val Glu Ala Leu Asp Lys Val Leu Leu Ala Val Ser Glu			
740	745	750	
Glu Tyr Leu Asp Leu Arg Leu Thr Phe Gly Pro Tyr Ser Pro Ser Gly			
755	760	765	
Gly Asp Ala Ser Ser Thr Cys Ser Ser Asp Ser Val Phe Ser His			
770	775	780	
Asp Pro Leu Pro Leu Gly Ser Ser Ser Phe Pro Phe Gly Ser Gly Val			
785	790	795	800

Gln Thr

<210> SEQ_ID NO 231
<211> LENGTH: 1014
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 231

Met Leu Ala Arg Ala Pro Pro Arg Arg Pro Pro Arg Leu Val Leu Leu			
1	5	10	15
Arg Leu Leu Leu Leu His Leu Leu Leu Leu Ala Leu Arg Ala Arg Cys			
20	25	30	
Leu Ser Ala Glu Pro Gly Gln Gly Ala Gln Thr Trp Ala Arg Phe Ala			
35	40	45	
Arg Ala Pro Ala Pro Glu Ala Ala Gly Leu Leu His Asp Thr Phe Pro			
50	55	60	
Asp Gly Phe Leu Trp Ala Val Gly Ser Ala Ala Tyr Gln Thr Glu Gly			
65	70	75	80
Gly Trp Arg Gln His Gly Lys Gly Ala Ser Ile Trp Asp Thr Phe Thr			
85	90	95	
His His Ser Gly Ala Ala Pro Ser Asp Ser Pro Ile Val Val Ala Pro			
100	105	110	
Ser Gly Ala Pro Ser Pro Pro Leu Ser Ser Thr Gly Asp Val Ala Ser			
115	120	125	
Asp Ser Tyr Asn Asn Val Tyr Arg Asp Thr Glu Gly Leu Arg Glu Leu			
130	135	140	
Gly Val Thr His Tyr Arg Phe Ser Ile Ser Trp Ala Arg Val Leu Pro			
145	150	155	160
Asn Gly Thr Ala Gly Thr Pro Asn Arg Glu Gly Leu Arg Tyr Tyr Arg			
165	170	175	
Arg Leu Leu Glu Arg Leu Arg Glu Leu Gly Val Gln Pro Val Val Thr			
180	185	190	
Leu Tyr His Trp Asp Leu Pro Gln Arg Leu Gln Asp Thr Tyr Gly Gly			

US 9,475,856 B2

395**396**

-continued

195 200 205

Trp Ala Asn Arg Ala Leu Ala Asp His Phe Arg Asp Tyr Ala Glu Leu
 210 215 220

Cys Phe Arg His Phe Gly Gly Gln Val Lys Tyr Trp Ile Thr Ile Asp
 225 230 235 240

Asn Pro Tyr Val Val Ala Trp His Gly Tyr Ala Thr Gly Arg Leu Ala
 245 250 255

Pro Gly Val Arg Gly Ser Ser Arg Leu Gly Tyr Leu Val Ala His Asn
 260 265 270

Leu Leu Leu Ala His Ala Lys Val Trp His Leu Tyr Asn Thr Ser Phe
 275 280 285

Arg Pro Thr Gln Gly Gly Arg Val Ser Ile Ala Leu Ser Ser His Trp
 290 295 300

Ile Asn Pro Arg Arg Met Thr Asp Tyr Asn Ile Arg Glu Cys Gln Lys
 305 310 315 320

Ser Leu Asp Phe Val Leu Gly Trp Phe Ala Lys Pro Ile Phe Ile Asp
 325 330 335

Gly Asp Tyr Pro Glu Ser Met Lys Asn Asn Leu Ser Ser Leu Leu Pro
 340 345 350

Asp Phe Thr Glu Ser Glu Lys Arg Leu Ile Arg Gly Thr Ala Asp Phe
 355 360 365

Phe Ala Leu Ser Phe Gly Pro Thr Leu Ser Phe Gln Leu Leu Asp Pro
 370 375 380

Asn Met Lys Phe Arg Gln Leu Glu Ser Pro Asn Leu Arg Gln Leu Leu
 385 390 395 400

Ser Trp Ile Asp Leu Glu Tyr Asn His Pro Pro Ile Phe Ile Val Glu
 405 410 415

Asn Gly Trp Phe Val Ser Gly Thr Thr Lys Arg Asp Asp Ala Lys Tyr
 420 425 430

Met Tyr Tyr Leu Lys Lys Phe Ile Met Glu Thr Leu Lys Ala Ile Arg
 435 440 445

Leu Asp Gly Val Asp Val Ile Gly Tyr Thr Ala Trp Ser Leu Met Asp
 450 455 460

Gly Phe Glu Trp His Arg Gly Tyr Ser Ile Arg Arg Gly Leu Phe Tyr
 465 470 475 480

Val Asp Phe Leu Ser Gln Asp Lys Glu Leu Leu Pro Lys Ser Ser Ala
 485 490 495

Leu Phe Tyr Gln Lys Leu Ile Glu Asp Asn Gly Phe Pro Pro Leu Pro
 500 505 510

Glu Asn Gln Pro Leu Glu Gly Thr Phe Pro Cys Asp Phe Ala Trp Gly
 515 520 525

Val Val Asp Asn Tyr Val Gln Val Asp Thr Thr Leu Ser Gln Phe Thr
 530 535 540

Asp Pro Asn Val Tyr Leu Trp Asp Val His His Ser Lys Arg Leu Ile
 545 550 555 560

Lys Val Asp Gly Val Val Ala Lys Lys Arg Lys Pro Tyr Cys Val Asp
 565 570 575

Phe Ser Ala Ile Arg Pro Gln Ile Thr Leu Leu Arg Glu Met Arg Val
 580 585 590

Thr His Phe Arg Phe Ser Leu Asp Trp Ala Leu Ile Leu Pro Leu Gly
 595 600 605

Asn Gln Thr Gln Val Asn His Thr Val Leu His Phe Tyr Arg Cys Met
 610 615 620

-continued

Ile Ser Glu Leu Val His Ala Asn Ile Thr Pro Val Val Ala Leu Trp
625 630 635 640

Gln Pro Ala Ala Pro His Gln Gly Leu Pro His Ala Leu Ala Lys His
645 650 655

Gly Ala Trp Glu Asn Pro His Thr Ala Leu Ala Phe Ala Asp Tyr Ala
660 665 670

Asn Leu Cys Phe Lys Glu Leu Gly His Trp Val Asn Leu Trp Ile Thr
675 680 685

Met Asn Glu Pro Asn Thr Arg Asn Met Thr Tyr Arg Ala Gly His His
690 695 700

Leu Leu Arg Ala His Ala Leu Ala Trp His Leu Tyr Asp Asp Lys Phe
705 710 715 720

Arg Ala Ala Gln Lys Gly Lys Ile Ser Ile Ala Leu Gln Ala Asp Trp
725 730 735

Ile Glu Pro Ala Cys Pro Phe Ser Gln Asn Asp Lys Glu Val Ala Glu
740 745 750

Arg Val Leu Glu Phe Asp Ile Gly Trp Leu Ala Glu Pro Ile Phe Gly
755 760 765

Ser Gly Asp Tyr Pro Arg Val Met Arg Asp Trp Leu Asn Gln Lys Asn
770 775 780

Asn Phe Leu Leu Pro Tyr Phe Thr Glu Asp Glu Lys Lys Leu Val Arg
785 790 795 800

Gly Ser Phe Asp Phe Leu Ala Val Ser His Tyr Thr Thr Ile Leu Val
805 810 815

Asp Trp Glu Lys Glu Asp Pro Met Lys Tyr Asn Asp Tyr Leu Glu Val
820 825 830

Gln Glu Met Thr Asp Ile Thr Trp Leu Asn Ser Pro Ser Gln Val Ala
835 840 845

Val Val Pro Trp Gly Leu Arg Lys Val Leu Asn Trp Leu Arg Phe Lys
850 855 860

Tyr Gly Asp Leu Pro Met Tyr Val Thr Ala Asn Gly Ile Asp Asp Asp
865 870 875 880

Pro His Ala Glu Gln Asp Ser Leu Arg Ile Tyr Tyr Ile Lys Asn Tyr
885 890 895

Val Asn Glu Ala Leu Lys Ala Tyr Val Leu Asp Asp Ile Asn Leu Cys
900 905 910

Gly Tyr Phe Ala Tyr Ser Leu Ser Asp Arg Ser Ala Pro Lys Ser Gly
915 920 925

Phe Tyr Arg Tyr Ala Ala Asn Gln Phe Glu Pro Lys Pro Ser Met Lys
930 935 940

His Tyr Arg Lys Ile Ile Asp Ser Asn Gly Phe Leu Gly Ser Gly Thr
945 950 955 960

Leu Gly Arg Phe Cys Pro Glu Glu Tyr Thr Val Cys Thr Glu Cys Gly
965 970 975

Phe Phe Gln Thr Arg Lys Ser Leu Leu Val Phe Ile Ser Phe Leu Val
980 985 990

Phe Thr Phe Ile Ile Ser Leu Ala Leu Ile Phe His Tyr Ser Lys Lys
995 1000 1005

Gly Gln Arg Ser Tyr Lys
1010

US 9,475,856 B2

399**400**

-continued

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 232

Met	Asp	Ser	Asp	Glu	Thr	Gly	Phe	Glu	His	Ser	Gly	Leu	Trp	Val	Ser
1				5				10					15		
Val	Leu	Ala	Gly	Leu	Leu	Leu	Gly	Ala	Cys	Gln	Ala	His	Pro	Ile	Pro
	20				25								30		
Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	Tyr
	35					40						45			
Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His	Leu	Glu	Ile	Arg
	50					55						60			
Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser	Pro	Glu	Ser	Leu
	65				70					75			80		
Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	Ile	Leu	Gly	Val
		85				90							95		
Lys	Thr	Ser	Arg	Phe	Leu	Cys	Gln	Arg	Pro	Asp	Gly	Ala	Leu	Tyr	Gly
		100				105							110		
Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg	Glu	Leu	Leu	Leu
	115				120							125			
Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	Gly	Leu	Pro	Leu
	130				135							140			
His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro	Ala	Pro	Arg	Gly
	145				150					155			160		
Pro	Ala	Arg	Phe	Leu	Pro	Leu	Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu
		165				170						175			
Pro	Pro	Gly	Ile	Leu	Ala	Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp
		180				185						190			
Pro	Leu	Ser	Met	Val	Gly	Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Phe	Ala
		195			200							205			

Ser

<210> SEQ_ID NO 233

<211> LENGTH: 209

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 233

Met	Asp	Ser	Asp	Glu	Thr	Gly	Phe	Glu	His	Ser	Gly	Leu	Trp	Val	Ser
1				5				10					15		
Val	Leu	Ala	Gly	Leu	Leu	Leu	Gly	Ala	Cys	Gln	Ala	His	Pro	Ile	Pro
	20				25								30		
Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val	Arg	Gln	Arg	Tyr
	35					40						45			
Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His	Leu	Glu	Ile	Arg
	50					55						60			
Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser	Pro	Glu	Ser	Leu
	65				70					75			80		
Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	Ile	Leu	Gly	Val
		85				90						95			
Lys	Thr	Ser	Arg	Phe	Leu	Cys	Gln	Arg	Pro	Asp	Gly	Ala	Leu	Tyr	Gly
		100				105						110			
Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg	Glu	Leu	Leu	Leu
	115				120							125			
Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	Gly	Leu	Pro	Leu

401**402**

-continued

130 135 140

His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
 145 150 155 160

Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
 165 170 175

Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
 180 185 190

Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Phe Glu
 195 200 205

Lys

<210> SEQ ID NO 234

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 234

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Lys
 35 40

<210> SEQ ID NO 235

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 235

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Glu Ser
 35 40

<210> SEQ ID NO 236

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 236

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Phe Ala Ser
 35 40

<210> SEQ ID NO 237

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

-continued

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 237

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1					5			10				15			
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
			20					25				30			
Ser	Gln	Val	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
					35			40							

<210> SEQ ID NO 238

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 238

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1					5			10				15			
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
			20					25				30			
Ser	Ala	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
					35			40							

<210> SEQ ID NO 239

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 239

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1					5			10				15			
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
			20					25				30			
Glu	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
					35			40							

<210> SEQ ID NO 240

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 240

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1					5			10				15			
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Leu
			20					25				30			
Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
					35			40							

<210> SEQ ID NO 241

<211> LENGTH: 43

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 241

-continued

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15
 Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Thr Gly
 20 25 30
 Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 242
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 242
 Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15
 Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Leu Val Gly Pro
 20 25 30
 Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 243
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 243
 Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15
 Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Gly Met Val Gly Pro
 20 25 30
 Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 244
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 244
 Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15
 Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Phe Ser Met Val Gly Pro
 20 25 30
 Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 245
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 245
 Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

-continued

Gln Pro Pro Asp Val Gly Ser Met Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 246
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 246

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Asp Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 247
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 247

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Thr Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 248
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 248

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Glu Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 249
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 249

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Leu Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

-continued

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 250
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 250

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Ser Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 251
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 251

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Ser
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 252
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 252

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Phe Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

<210> SEQ ID NO 253
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 253

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Met Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
 20 25 30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
 35 40

-continued

```
<210> SEQ_ID NO 254
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER_INFORMATION: FGF21 Variant

<400> SEQUENCE: 254
```

```

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Asp Leu Ala Pro
1           5           10          15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
35          40

```

```
<210> SEQ ID NO 255
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant
```

```

<400> SEQUENCE: 235

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Ser Ile Leu Ala Pro
1           5           10          15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
25          30

```

```
<210> SEQ_ID NO 256
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER_INFORMATION: FGF21 Variant

>400> SEQUENCE: 256
```

```

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Glu Gly Ile Leu Ala Pro
1           5           10          15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
35          40

```

```
<210> SEQ_ID NO 257
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER_INFORMATION: FGF21 Variant
```

```

<400> SEQUENCE: 257

Pro Gly Leu Pro Pro Ala Leu Pro Glu Leu Pro Gly Ile Leu Ala Pro
1           5           10          15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser

```

1210: SEQ ID NO 252

-continued

<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 258

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	His	Pro	Pro	Gly	Ile	Leu	Ala
1															
															15
Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly
															30
Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser					

35 40

<210> SEQ ID NO 259
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 259

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Gly	Pro	Pro	Gly	Ile	Leu	Ala
1															
															15
Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly
															30
Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser					

35 40

<210> SEQ ID NO 260
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 260

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Arg	Pro	Pro	Gly	Ile	Leu	Ala
1															
															15
Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly
															30

Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
35 40

<210> SEQ ID NO 261
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 261

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Ile	Leu	Ala				
1															
															15
Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly
															30
Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser					

35 40

<210> SEQ ID NO 262
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial

-continued

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 262

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Asp	Pro	Pro	Gly	Ile	Leu	Ala
1							5			10				15	
Pro	Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly
	20						25						30		
Pro	Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser					
	35						40								

<210> SEQ_ID NO 263

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 263

Pro	Gly	Leu	Pro	Pro	Ala	Glu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1							5			10			15		
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
	20						25					30			
Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
	35						40								

<210> SEQ_ID NO 264

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 264

Pro	Gly	Leu	Pro	Pro	Glu	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1							5			10			15		
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
	20						25					30			
Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
	35						40								

<210> SEQ_ID NO 265

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 265

Pro	Gly	Leu	Val	Pro	Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro
1							5			10			15		
Gln	Pro	Pro	Asp	Val	Gly	Ser	Ser	Asp	Pro	Leu	Ser	Met	Val	Gly	Pro
	20						25					30			
Ser	Gln	Gly	Arg	Ser	Pro	Ser	Tyr	Ala	Ser						
	35						40								

<210> SEQ_ID NO 266

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

-continued

<400> SEQUENCE: 266

```

Pro Gly Met Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5           10           15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
35          40

```

<210> SEQ ID NO 267

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 267

```

Pro Pro Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5           10           15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
35          40

```

<210> SEQ ID NO 268

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 268

```

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5           10           15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Tyr Glu Lys
35          40

```

<210> SEQ ID NO 269

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 269

```

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5           10           15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro
20          25          30

Ser Gln Gly Arg Ser Pro Ser Phe Glu Lys
35          40

```

<210> SEQ ID NO 270

<211> LENGTH: 42

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 270

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro

419**420**

-continued

1	5	10	15
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro			
20	25	30	
Ser Gln Val Arg Ser Pro Ser Phe Glu Lys			
35	40		

<210> SEQ ID NO 271
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 271

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro			
1	5	10	15
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro			
20	25	30	
Ser Ala Val Arg Ser Pro Ser Phe Glu Lys			
35	40		

<210> SEQ ID NO 272
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 272

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro			
1	5	10	15
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro			
20	25	30	
Glu Ala Val Arg Ser Pro Ser Phe Glu Lys			
35	40		

<210> SEQ ID NO 273
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 273

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro			
1	5	10	15
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Leu			
20	25	30	
Glu Ala Val Arg Ser Pro Ser Phe Glu Lys			
35	40		

<210> SEQ ID NO 274
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 274

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro			
1	5	10	15
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Thr Gly			

US 9,475,856 B2

421**422**

-continued

20

25

30

Leu Glu Ala Val Arg Ser Pro Phe Glu Lys
 35 40

<210> SEQ ID NO 275
 <211> LENGTH: 43
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 275

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Leu Val Thr Gly
 20 25 30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40

<210> SEQ ID NO 276
 <211> LENGTH: 43
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 276

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Gly Leu Val Thr Gly
 20 25 30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40

<210> SEQ ID NO 277
 <211> LENGTH: 43
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 277

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Phe Gly Leu Val Thr Gly
 20 25 30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40

<210> SEQ ID NO 278
 <211> LENGTH: 43
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 278

Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
 1 5 10 15

Gln Pro Pro Asp Val Gly Ser Met Asp Pro Phe Gly Leu Val Thr Gly
 20 25 30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys

US 9,475,856 B2

423**424**

-continued

35

40

<210> SEQ ID NO 279
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 279

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5          10          15

Gln Pro Pro Asp Val Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 280
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 280

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5          10          15

Gln Pro Pro Asp Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 281
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 281

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5          10          15

Gln Pro Pro Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 282
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 282

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5          10          15

Gln Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

-continued

<210> SEQ ID NO 283
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 283

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
1           5          10          15

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 284
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 284

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Ser
1           5          10          15

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 285
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 285

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Phe Ser
1           5          10          15

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 286
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 286

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Met Phe Ser
1           5          10          15

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 287
<211> LENGTH: 43
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 287

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Asp Met Phe Ser
1           5          10          15
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 288
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 288

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Ser Asp Met Phe Ser
1           5          10          15
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 289
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 289

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Glu Ser Asp Met Phe Ser
1           5          10          15
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 290
<211> LENGTH: 43
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 290

```
Pro Gly Leu Pro Pro Ala Leu Pro Glu Leu Glu Ser Asp Met Phe Ser
1           5          10          15
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
20          25          30
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35          40
```

<210> SEQ ID NO 291
<211> LENGTH: 44
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

-continued

<400> SEQUENCE: 291

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	His	Leu	Glu	Ser	Asp	Met	Phe
1															
															15
Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr
															30
Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys				
															40
															35

<210> SEQ ID NO 292

<211> LENGTH: 45

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 292

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Gly	His	Leu	Glu	Ser	Asp	Met
1															15
Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val
															30
20															25
Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys			
															45
															35

<210> SEQ ID NO 293

<211> LENGTH: 46

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 293

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Arg	Gly	His	Leu	Glu	Ser	Asp
1															15
Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu
															30
20															25
Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys		
															45
															35

<210> SEQ ID NO 294

<211> LENGTH: 47

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 294

Pro	Gly	Leu	Pro	Pro	Ala	Leu	Pro	Glu	Leu	Arg	Gly	His	Leu	Glu	Ser
1															15
Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly
															30
20															25
Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys	
															45
															35

<210> SEQ ID NO 295

<211> LENGTH: 48

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 295

US 9,475,856 B2

431**432**

-continued

Pro Gly Leu Pro Pro Ala Leu Pro Glu Asp Leu Arg Gly His Leu Glu
 1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
 20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40 45

<210> SEQ ID NO 296
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 296

Pro Gly Leu Pro Pro Ala Glu Pro Glu Asp Leu Arg Gly His Leu Glu
 1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
 20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40 45

<210> SEQ ID NO 297
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 297

Pro Gly Leu Pro Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
 1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
 20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40 45

<210> SEQ ID NO 298
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 298

Pro Gly Leu Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
 1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
 20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 35 40 45

<210> SEQ ID NO 299
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 299

Pro Gly Met Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
 1 5 10 15

-continued

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35 40 45

<210> SEQ ID NO 300
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF21 Variant

<400> SEQUENCE: 300

Pro Pro Met Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
35 40 45

<210> SEQ ID NO 301
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 301

Leu Pro Met Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20 25 30

Gly Leu Val Thr Gly Leu Glu Ala Val Arg Ser Pro Ser Phe Ala Lys
35 40 45

<210> SEQ ID NO 302
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 302

Leu Pro Met Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20 25 30

Gly Leu Val Thr Gly Leu Ser Ala Val Arg Ser Pro Ser Phe Glu Lys
35 40 45

<210> SEQ ID NO 303
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 303

Leu Pro Met Val Pro Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu
1 5 10 15

Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe
20 25 30

-continued

<210> SEQ ID NO 308
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 308

Leu	Pro	Met	Val	Pro	Ala	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu
1															
															15
Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe
															30
Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys
															45
35															
40															

<210> SEQ ID NO 309
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 309

Leu	Pro	Met	Pro	Pro	Glu	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu
1															15
Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe
															30
Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys
															45
35															
40															

<210> SEQ ID NO 310
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 310

Leu	Pro	Leu	Val	Pro	Glu	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu
1															15
Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe
															30
Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys
															45
35															
40															

<210> SEQ ID NO 311
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: FGF19 Variant

<400> SEQUENCE: 311

Leu	Gly	Met	Val	Pro	Glu	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu
1															15
Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe
															30
Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys
															45
35															
40															

<210> SEQ ID NO 312
<211> LENGTH: 187

US 9,475,856 B2

439**440**

-continued

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 312

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val	1	5	10	15
---	---	---	----	----

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His	20	25	30
---	----	----	----

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser	35	40	45
---	----	----	----

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln	50	55	60
---	----	----	----

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly	65	70	75	80
---	----	----	----	----

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg	85	90	95
---	----	----	----

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His	100	105	110
---	-----	-----	-----

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro	115	120	125
---	-----	-----	-----

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Pro Met Val Pro	130	135	140
---	-----	-----	-----

Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser	145	150	155	160
---	-----	-----	-----	-----

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly	165	170	175
---	-----	-----	-----

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys	180	185
---	-----	-----

<210> SEQ ID NO 313

<211> LENGTH: 187

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 313

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val	1	5	10	15
---	---	---	----	----

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His	20	25	30
---	----	----	----

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser	35	40	45
---	----	----	----

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln	50	55	60
---	----	----	----

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly	65	70	75	80
---	----	----	----	----

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg	85	90	95
---	----	----	----

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His	100	105	110
---	-----	-----	-----

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro	115	120	125
---	-----	-----	-----

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Met Val Pro	130	135	140
---	-----	-----	-----

US 9,475,856 B2

441

442

-continued

Glu	Glu	Pro	Glu	Asp	Leu	Arg	Gly	His	Leu	Glu	Ser	Asp	Met	Phe	Ser
145					150					155					160
Ser	Pro	Leu	Glu	Thr	Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr	Gly
						165			170					175	
Leu	Glu	Ala	Val	Arg	Ser	Pro	Ser	Phe	Glu	Lys					
						180			185						

<210> SEQ ID NO 314
<211> LENGTH: 187
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE : 314

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Val Pro
130 135 140

Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser
 145 150 155 160

Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
165 170 175

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
180 185

<210> SEQ ID NO 315
<211> LENGTH: 187
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 315

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly

-continued

65	70	75	80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg			
85	90	95	
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His			
100	105	110	
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro			
115	120	125	
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro			
130	135	140	
Glu Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser			
145	150	155	160
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly			
165	170	175	
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys			
180	185		

<210> SEQ ID NO 316

<211> LENGTH: 187

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 316

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val			
1	5	10	15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His			
20	25	30	
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser			
35	40	45	
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln			
50	55	60	
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly			
65	70	75	80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg			
85	90	95	
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His			
100	105	110	
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro			
115	120	125	
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro			
130	135	140	
Ala Glu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser			
145	150	155	160
Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly			
165	170	175	
Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys			
180	185		

<210> SEQ ID NO 317

<211> LENGTH: 187

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 317

US 9,475,856 B2

445**446**

-continued

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15
 Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30
 Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45
 Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60
 Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80
 Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95
 Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110
 Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125
 Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140
 Ala Leu Pro Glu Asp Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser
 145 150 155 160
 Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
 165 170 175
 Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
 180 185

<210> SEQ ID NO 318
 <211> LENGTH: 187
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 318

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15
 Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30
 Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45
 Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60
 Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80
 Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95
 Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110
 Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125
 Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140
 Ala Leu Pro Glu Pro Leu Arg Gly His Leu Glu Ser Asp Met Phe Ser
 145 150 155 160
 Ser Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly
 165 170 175

US 9,475,856 B2

447**448**

-continued

Leu Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
180 185

<210> SEQ ID NO 319
<211> LENGTH: 186
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 319

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Arg Gly His Leu Glu Ser Asp Met Phe Ser Ser
145 150 155 160

Pro Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu
165 170 175

Glu Ala Val Arg Ser Pro Ser Phe Glu Lys
180 185

<210> SEQ ID NO 320
<211> LENGTH: 185
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 320

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

-continued

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140
Ala Leu Pro Glu Pro Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro
145 150 155 160
Leu Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu
165 170 175
Ala Val Arg Ser Pro Ser Phe Glu Lys
180 185

<210> SEQ ID NO 321
<211> LENGTH: 184
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 321

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140
Ala Leu Pro Glu Pro His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu
145 150 155 160
Glu Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala
165 170 175
Val Arg Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 322
<211> LENGTH: 183
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 322

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser

US 9,475,856 B2

451**452**

-continued

35	40	45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln		
50	55	60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly		
65	70	75
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg		
85	90	95
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His		
100	105	110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro		
115	120	125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro		
130	135	140
Ala Leu Pro Glu Pro Leu Glu Ser Asp Met Phe Ser Ser Pro Leu Glu		
145	150	155
Thr Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val		
165	170	175
Arg Ser Pro Ser Phe Glu Lys		
180		

<210> SEQ ID NO 323
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 323

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val		
1	5	10
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His		
20	25	30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser		
35	40	45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln		
50	55	60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly		
65	70	75
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg		
85	90	95
Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His		
100	105	110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro		
115	120	125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro		
130	135	140
Ala Leu Pro Glu Pro Glu Ser Asp Met Phe Ser Ser Pro Leu Glu Thr		
145	150	155
Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg		
165	170	175
Ser Pro Ser Phe Glu Lys		
180		

<210> SEQ ID NO 324
<211> LENGTH: 182
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 324

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val
1								5	10					15	

Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His
			20			25						30			

Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser
							35		40			45			

Pro	Glu	Ser	Leu	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln
							50		55			60			

Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Met	Arg	Pro	Asp	Gly
65							70			75		80			

Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg
							85		90			95			

Glu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His
						100		105			110			

Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro
							115		120			125			

Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Gly	Leu	Pro	Pro		
							130		135			140			

Ala	Leu	Pro	Glu	Pro	Pro	Ser	Asp	Met	Phe	Ser	Ser	Pro	Leu	Glu	Thr
145							150		155			160			

Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg
						165		170			175				

Ser	Pro	Ser	Phe	Glu	Lys										
					180										

<210> SEQ_ID NO 325

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 325

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val
1								5	10					15	

Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His
			20			25						30			

Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser
							35		40			45			

Pro	Glu	Ser	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln	
							50		55			60			

Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Met	Arg	Pro	Asp	Gly
65							70			75		80			

Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg
							85		90			95			

Glu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His
						100		105			110			

Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro
							115		120			125			

Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Gly	Leu	Pro	Pro		
							130		135			140			

US 9,475,856 B2

455**456**

-continued

Ala Leu Pro Glu Pro Pro Gly Asp Met Phe Ser Ser Pro Leu Glu Thr
 145 150 155 160
 Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
 165 170 175
 Ser Pro Ser Phe Glu Lys
 180

<210> SEQ ID NO 326
 <211> LENGTH: 182
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 326

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15
 Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30
 Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45
 Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60
 Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80
 Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95
 Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110
 Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125
 Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140
 Ala Leu Pro Glu Pro Pro Gly Ile Met Phe Ser Ser Pro Leu Glu Thr
 145 150 155 160
 Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
 165 170 175
 Ser Pro Ser Phe Glu Lys
 180

<210> SEQ ID NO 327
 <211> LENGTH: 182
 <212> TYPE: PRT
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 327

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15
 Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30
 Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45
 Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60
 Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80

-continued

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Phe Ser Ser Pro Leu Glu Thr
145 150 155 160

Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 328
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 328

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Ser Ser Pro Leu Glu Thr
145 150 155 160

Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 329
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 329

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val

US 9,475,856 B2

459**460**

-continued

1	5	10	15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His			
20	25	30	
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser			
35	40	45	
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln			
50	55	60	
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly			
65	70	75	80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg			
85	90	95	
Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His			
100	105	110	
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro			
115	120	125	
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro			
130	135	140	
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Ser Pro Leu Glu Thr			
145	150	155	160
Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg			
165	170	175	
Ser Pro Ser Phe Glu Lys			
180			

<210> SEQ ID NO 330

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 330

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val			
1	5	10	15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His			
20	25	30	
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser			
35	40	45	
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln			
50	55	60	
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly			
65	70	75	80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg			
85	90	95	
Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His			
100	105	110	
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro			
115	120	125	
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro			
130	135	140	
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Leu Glu Thr			
145	150	155	160
Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg			
165	170	175	
Ser Pro Ser Phe Glu Lys			

-continued

180

<210> SEQ ID NO 331
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 331

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val
1				5				10			15				

Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His
	20				25							30			

Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser
	35				40					45					

Pro	Glu	Ser	Leu	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln
	50				55				60						

Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Met	Arg	Pro	Asp	Gly
65				70				75			80				

Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg
	85					90					95				

Glu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	
	100				105				110						

Gly	Leu	Pro	Leu	His	Leu	Pro	Gly	Asn	Lys	Ser	Pro	His	Arg	Asp	Pro
	115				120				125						

Ala	Pro	Arg	Gly	Pro	Ala	Arg	Phe	Leu	Pro	Gly	Leu	Pro	Pro		
	130				135			140							

Ala	Leu	Pro	Glu	Pro	Pro	Gly	Ile	Leu	Ala	Pro	Gln	Pro	Pro	Glu	Thr
145					150			155			160				

Asp	Ser	Met	Asp	Pro	Phe	Gly	Leu	Val	Thr	Gly	Leu	Glu	Ala	Val	Arg
	165				170				175						

Ser	Pro	Ser	Phe	Glu	Lys										
	180														

<210> SEQ ID NO 332
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 332

His	Pro	Ile	Pro	Asp	Ser	Ser	Pro	Leu	Leu	Gln	Phe	Gly	Gly	Gln	Val
1				5				10			15				

Arg	Gln	Arg	Tyr	Leu	Tyr	Thr	Asp	Asp	Ala	Gln	Gln	Thr	Glu	Ala	His
	20				25							30			

Leu	Glu	Ile	Arg	Glu	Asp	Gly	Thr	Val	Gly	Gly	Ala	Ala	Asp	Gln	Ser
	35				40				45						

Pro	Glu	Ser	Leu	Leu	Gln	Leu	Lys	Ala	Leu	Lys	Pro	Gly	Val	Ile	Gln
	50				55				60						

Ile	Leu	Gly	Val	Lys	Thr	Ser	Arg	Phe	Leu	Cys	Met	Arg	Pro	Asp	Gly
65				70				75			80				

Ala	Leu	Tyr	Gly	Ser	Leu	His	Phe	Asp	Pro	Glu	Ala	Cys	Ser	Phe	Arg
	85					90					95				

Glu	Leu	Leu	Glu	Asp	Gly	Tyr	Asn	Val	Tyr	Gln	Ser	Glu	Ala	His	
	100				105				110						

-continued

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Thr
145 150 155 160

Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 333
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 333

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95

Glu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145 150 155 160

Asp Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
165 170 175

Ser Pro Ser Phe Glu Lys
180

<210> SEQ ID NO 334
<211> LENGTH: 182
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 334

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35 40 45

-continued

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
 145 150 155 160

Gly Ser Ser Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
 165 170 175

Ser Pro Ser Phe Glu Lys
 180

<210> SEQ ID NO 335

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 335

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1 5 10 15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
 20 25 30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
 35 40 45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
 50 55 60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
 65 70 75 80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
 85 90 95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
 100 105 110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
 115 120 125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
 130 135 140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
 145 150 155 160

Gly Ser Ser Asp Pro Leu Gly Leu Val Thr Gly Leu Glu Ala Val Arg
 165 170 175

Ser Pro Ser Phe Glu Lys
 180

<210> SEQ ID NO 336

<211> LENGTH: 182

<212> TYPE: PRT

<213> ORGANISM: Artificial

-continued

<220> FEATURE:

<223> OTHER INFORMATION: Chimeric Protein

<400> SEQUENCE: 336

```

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
1           5          10          15

Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20          25          30

Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
35          40          45

Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50          55          60

Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Met Arg Pro Asp Gly
65          70          75          80

Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85          90          95

Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100         105         110

Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115         120         125

Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130         135         140

Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145         150         155         160

Gly Ser Ser Asp Pro Leu Ser Leu Val Thr Gly Leu Glu Ala Val Arg
165         170         175

Ser Pro Ser Phe Glu Lys
180

```

What is claimed:

1. An isolated chimeric protein comprising:
an N-terminus coupled to a C-terminus, wherein the
N-terminus comprises an N-terminal portion of fibroblast growth factor (FGF) 21, and the C-terminus
comprises a C-terminal portion of FGF19, wherein the
N-terminal portion of FGF21 comprises an amino acid
sequence beginning at residue 29 and ending at any one of residues 167-197 of SEQ ID NO:152, wherein the
C-terminal portion of FGF21 beginning at any one of residues 167-197 and ending at residue 209 of SEQ ID
NO: 152 is progressively replaced with a C-terminal portion of FGF19 beginning at any one of residues
169-204 and ending at residue 216 of SEQ ID NO:1. 40
 2. The chimeric protein of claim 1, wherein the chimeric protein comprises the amino acid sequence of SEQ ID NO: 208, SEQ ID NO: 209, or SEQ ID NO: 210. 50
 3. A pharmaceutical composition comprising:
the chimeric protein of claim 1 and
a pharmaceutically-acceptable carrier. 55
 4. The pharmaceutical composition of claim 3 further comprising:
an anti-inflammatory agent, an antifibrotic agent, an anti-hypertensive agent, an antidiabetic agent, a triglyceride-lowering agent, and/or a cholesterol-lowering agent. 60
 5. The chimeric protein according to claim 1, wherein the chimeric protein consists of the amino acid sequence of SEQ ID NO: 208, SEQ ID NO: 209, or SEQ ID NO: 210. 65
 6. The chimeric protein according to claim 1, wherein the chimeric protein comprises the amino acid sequence of SEQ
- ID NO:312, SEQ ID NO:313, SEQ ID NO:314, SEQ ID NO:315, SEQ ID NO:316, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:319, SEQ ID NO:320, SEQ ID NO:321, SEQ ID NO:322, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:325, SEQ ID NO:326, SEQ ID NO:327, SEQ ID NO:328, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:331, SEQ ID NO:332, SEQ ID NO:333, SEQ ID NO:334, SEQ ID NO:335, or SEQ ID NO:336.
7. A method for decreasing blood glucose level in a subject, said method comprising:
selecting a subject and
administering to said selected subject an effective amount
of the chimeric protein of claim 1 to decrease blood
glucose level in the selected subject. 70
 8. The method according to claim 7, wherein the selected subject has diabetes, obesity, or metabolic syndrome. 75
 9. The method according to claim 8, wherein the chimeric protein is administered with a pharmaceutically-acceptable carrier. 80
 10. The method according to claim 8, wherein the selected subject is a mammal. 85
 11. The method according to claim 10, wherein the selected subject is a human. 90
 12. The method according to claim 8, wherein the selected subject has diabetes. 95
 13. The method according to claim 12, wherein the selected subject has type II diabetes, gestational diabetes, or drug-induced diabetes. 100
 14. The method according to claim 12, wherein the selected subject has type I diabetes. 105

469

- 15.** The method according to claim **8**, wherein the selected subject has obesity.
16. The method according to claim **8**, wherein the selected subject has metabolic syndrome.
17. The method according to claim **7**, wherein said administering is carried out parenterally, subcutaneously, intravenously, intramuscularly, intraperitoneally, by intranasal instillation, by implantation, by intracavitary or intravesical instillation, intraocularly, intraarterially, intralesionally, transdermally, or by application to mucous membranes. ⁵ ¹⁰

18. The method according to claim **7**, wherein the chimeric protein comprises the amino acid sequence of SEQ ID NO: 208, SEQ ID NO: 209, or SEQ ID NO: 210.

19. The method according to claim **7**, wherein the chimeric protein consists of the amino acid sequence of SEQ ID ¹⁵ NO: 208, SEQ ID NO: 209, or SEQ ID NO: 210.

20. The method according to claim **7**, wherein the chimeric protein comprises the amino acid sequence of SEQ ID NO:312, SEQ ID NO:313, SEQ ID NO:314, SEQ ID NO:315, SEQ ID NO:316, SEQ ID NO:317, SEQ ID ²⁰ NO:318, SEQ ID NO:319, SEQ ID NO:320, SEQ ID NO:321, SEQ ID NO:322, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:325, SEQ ID NO:326, SEQ ID NO:327, SEQ ID NO:328, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:331, SEQ ID NO:332, SEQ ID ²⁵ NO:333, SEQ ID NO:334, SEQ ID NO:335, or SEQ ID NO:336.

* * * * *

470